

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2005, 05:38:23 ; Search time 21022 Seconds
(without alignments)
17642.290 Million cell updates/sec

Title: US-10-006-771B-1
Perfect score: 7654
Sequence: 1 aagctgcagctgcaggt.....gtaaaacagcgccagtgcc 7654

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3666.6	47.9	6501	12 AF132211	AF132211 Cloning v
2	3349	43.8	9320	6 AX663075	AX663075 Sequence
3	3307.8	43.2	6644	6 E23356	E23356 Virus vecto
4	3244.4	42.4	6505	6 AX823827	AX823827 Sequence
5	3233.6	42.1	5903	6 C0879092	C0879092 Sequence
6	3213	42.0	5782	6 AX491314	AX491314 Sequence
7	3183.2	41.6	7372	6 E23357	E23357 Virus vecto
8	3179.2	41.5	6277	12 AB086385	AB086385 Retrovira
9	3140	41.0	6248	12 AB041928	AB041928 Retrovira
10	3130.2	40.9	7257	6 C0879091	C0879091 Sequence
11	3113	40.7	6076	12 AB086386	AB086386 Retrovira
12	3111	40.6	5364	6 AR028670	AR028670 Sequence
13	3104.6	40.6	6506	14 U00220	U00220 Human Immun
14	3096	40.4	6522	6 BD138563	BD138563 Expressio
15	3094.8	40.4	6374	12 SYNNOV1	M64753 Moloney mur
16	3094	40.4	6365	6 I70974	I70974 Sequence 1
17	3087.4	40.3	6824	12 AB086387	AB086387 Retrovira
18	3083.4	40.3	5292	6 AR028669	AR028669 Sequence
19	3083.4	40.3	5294	6 AR028668	AR028668 Sequence

```

20 3083.4 40.3 5323 6 AR028667
21 3082.2 40.3 6219 6 BD222120
22 3082.2 40.3 6219 6 AX146810
23 3062.2 40.0 6145 6 AR140313
24 3062.2 40.0 6145 6 I77211
25 3062.2 40.0 6145 6 AR533381
26 3062.2 40.0 6145 6 BD105951
27 3062.2 40.0 6145 12 SYNMMPLN6
28 3050.2 39.9 6620 6 AR302094
29 3007 39.3 6259 12 SYNRRV
30 3006.8 39.3 6620 6 AX823828
31 3005.2 39.3 6620 12 SYNMMPLN3
32 2980.6 38.9 6690 12 AB086389
33 2960.2 38.7 5901 12 AB086388
34 2959.8 38.5 5883 12 SYNMMPLN2
35 2946.6 38.5 5904 12 AB041927
36 2939 38.4 5904 12 AB086384
37 2899.2 37.9 6141 6 BD138562
38 2860.4 37.4 6094 12 AF113968
39 2842.4 37.1 6072 12 D88622
40 2835.2 37.0 6837 12 SYNNOV2
41 2783.8 36.4 7797 6 E23355
42 2703 35.3 5865 6 AR164477
43 2687.4 35.1 5874 6 BD138565
44 2687.4 35.1 5874 12 SYNMMPLN4
45 2675.4 35.0 6283 6 AX133941

```

ALIGNMENTS

```

RESULT 1
AF132211
LOCUS AF132211 6501 bp DNA linear SYN 18-MAY-1999
DEFINITION Cloning vector pFBneo, complete sequence.
ACCESSION AF132211
VERSION AF132211.1 GI:4838522
KEYWORDS
SOURCE
ORGANISM Cloning vector pFBneo
Cloning vector pFBneo
Other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6501)
AUTHORS Felts,K. and Grafsky,A.J.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-1999) Technical Services, Stratagene, 11011 N.
Torrey Pines Rd., La Jolla, CA 92037, USA.
FEATURES
source
1..6501
Location/Qualifiers
/organism="Cloning vector pFBneo"
/mol_type="genomic DNA"
/specific_host="Escherichia coli"
/db_xref="taxon:93464"
209..760
/notes="5'"
760..2046
/misc_feature
/notes="Psi+; extended viral packaging signal"
2057..2086
/misc_feature
/notes="multiple cloning site"
2092..2680
/misc_feature
/notes="IRES"
2678..3487
/gene
/genes="neo"
2678..3487
/gene="neo"
/function="neomycin resistance"
/codon_start=1
/transl_table=11
/product="aminoglycoside 3'-phosphotransferase"
/protein_id="AAD31025.1"
/db_xref="GI:4838524"
/translations="MGSFRMIEODGLHAGSPAAWVERLFGVDWAQOTIGCSDAAVFRL
SAQRPVLFVKTKDLSGALNEIQDEARLWLTATGTCVCAAVLDVVTTEAGRDWLLIGEV
PGQDLSSHLAPAEKVSINADAMRRRLHTLDPATCFPDQAKHRIERARTRMEAGLVDQ

```


QY 2345 AAAGTAGACGCATCGCAGCTTGGATACACCGCGCCACGTGAAGGCTGCCGACCCCGGG 2404
Db |||||
QY 1964 AAAGTAGACGCATCGCAGCTTGGATACACCGCGCCACGTGAAGGCTGCCGACCCCGGG 2023
Db |||||
QY 2405 GGTGACCATCTCTAGACTCCATCGGATGGAGCTGTATCATCTCTCTTCTTGTGTAGCAA 2464
Db |||||
QY 2024 GGTGACCATCTCTAGACTCCGATCGAATTTGTGACGAATTCGGATCTCTCG-AGCGG 2082
Db |||||
QY 2465 CAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAG----- 2513
Db |||||
QY 2083 CCGCGATCCGGTATTTTCCACCATATTTCGCGTCTTTTGGCAATGTGAGGSCCGGAAAC 2142
Db |||||
QY 2514 -----CTTGAGCGCCAGCGTGGTGAACAGATGACCATCATCTGTAAAGCCAGTCAAGAT 2568
Db |||||
QY 2143 CTGCGCCCTGTCTTCTGACGAGCATCTCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGC 2202
Db |||||
QY 2569 GTGGGTACTCTGTAGCTTGTGTACACGAGAACCCAGTAAAGGCTC-CAAGCTGCTGTAT 2627
Db |||||
QY 2203 AAGTCTGTGTAATGTGTGAAGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAA 2262
Db |||||
QY 2628 CTACTGACATPCCACCCGGACACATGTGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGG 2687
Db |||||
QY 2263 CGTCTGTAGCACCTTTTGAGGACGCGGAACCCCCACCTGGCGACAGGTGCTCTGCG 2322
Db |||||
QY 2688 TACCGACTTACCTTACCATCAGCAGCTCCAGCCAGAGGACATCGCCACTTACTACTG 2747
Db |||||
QY 2323 GCCAAAGGCCACGTGTATAGATACACCTGCAAAAGGGGCGACAAACCCAGTGCACGTTG 2382
Db |||||
QY 2748 CCAGCAATATAGCTCTATCGGTCTGTGCGGCAAGGACCAAGGTGGGAATCAACACAGG 2807
Db |||||
QY 2383 TGAAGTGGATGTGTGGAAGAGTCAAAAGGCTCTCTCAAGGATATTCAACAAAGGGC 2442
Db |||||
QY 2808 TGGCTCAGGATCGGGTGGATCCGGCTCTGTGGCTCAGGATCGGAGTCCAACTGTGTGA 2867
Db |||||
QY 2443 TGAAGGATGCCAAGAGGTACCCATTTGTATGGGATCTGATCTGGGCGCTCGGTGCAT 2502
Db |||||
QY 2868 GAGCGGTGGAGGTGTGTGCAACCTGGCGGTCTCTGCGCTGTCTGCTCC--GCATCT 2925
Db |||||
QY 2503 GCTTTACATGTGTTAGTTCGAGGTATAAAACGCTAGGCCCCCGAAGCACGCGGAGCT 2562
Db |||||
QY 2926 GGCTTCGATTTACCAACATATTGATGATGTGGTGTGACGAGCACCTGGAAAGGTCTT 2985
Db |||||
QY 2563 GGTTTTCTTTGAAAAACACGATGATATATGGAACAAAACTTATTTCTGAAGAAGACT 2622
Db |||||
QY 2986 GAGTGGATTGGAGAAATTCATCCAGATAGCAGTACGATTAACATATG- -GCCGTCTCTAA 3043
Db |||||
QY 2623 ---TGGACACCAACTTCTGCGCTCGATTTCTCCACCTGTGTGCCAGCGGTTCAT 2679
Db |||||
QY 3044 AGGATAGATTTACAAATATCGCGAGACAAACGCAAGAACACATTTGCTCTGCAATGGACA 3103
Db |||||
QY 2680 GGGATCGTTTCGCATGATTGAACAGATGGATTGACGCGAGTTCTCGGCGCGTTGGGT 2739
Db |||||
QY 3104 GCCTGAGACCCGAAGACACCGGGGTCTATTTTGTGCAAGCTTTTATTTGGCTTCCCT 3163
Db |||||
QY 2740 GGAGAGGCTATTCGGCTATGACTGGGCAACACAGACAAATCGGCTGCTCTGATGCGCGGT 2799
Db |||||
QY 3164 GGTTTGCTTATTGGGCCAAGGACCCCGGTCCAGCTCTCAGTGTCTAAGCCCAACACAGA 3223
Db |||||
QY 2800 GTTCCGGCTGTGACGCGAGGGGCGCCGGTTCTTTTGTCAAGACCCAGCTGTCCGGTGC 2859
Db |||||
QY 3224 CGCCAGCGCGCGACCAACACCGCGGCCACCATCGCTCGCAGCCCTGTGCCCTGTC 3283
Db |||||
QY 2860 CTTGAATG-----AACTGAGGACGAGGACGCGCGGTATCTGTGCTGCCACGACGCGG 2915
Db |||||
QY 3284 GCCCAGAGGGGCTCGGCCAGCGCGGGGGCGCAGTGCACACAGAGGGGCTGCACTTCG 3343
Db |||||
QY 2916 TTCCTTGGCAGCTGTGTCTGACCTGTGTCTGAAAGCGGGAAGGACTGTGGCTGTATTGG 2975
Db |||||
QY 3344 CCCTGGATCCCAACTCTGCTACTCTGTGGATGGAATCTCTTCACTATATGGTGTCAATC 3403
Db |||||
QY 2976 GCGAAGTGCCTGGGCGAGGATCTCTCTGTCACTCACTTGTCTCTGCCGAGAAAGTATCCA 3035
Db |||||

QY 3404 TCACCTGCTTGTCTCTGAGAGTGAAGTTTCAGCAGGAGCGCAGACCCCGCGGTACCAGC 3463
Db |||||
QY 3036 TCATGGCTGATGCAATCGCGCGCTGCATACGCTTGTATCCGGCTACCTGCGCATTCGACC 3095
Db |||||
QY 3464 AGGCCAGAACCAAGCTCTATAACGAGCTCAATCTAGGACGAAGAGAGGATACGATGTTT 3523
Db |||||
QY 3096 ACCAAGCGAAACATCGC-----ATCGAGCGACACGTACTCTCGATGGAAGCGGTC 3146
Db |||||
QY 3524 TGGCAAGAGACGTGGCCGGGACCTTGATGATGGGGGGAAGAACCGAGAAAGAACCCCTC 3583
Db |||||
QY 3147 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGG- -GCTCGCGCCAGCGCAACTGTT 3204
Db |||||
QY 3584 AGGAAGCCTGTCAATGAATGACAGAAAGATAGATGGCGGAGCCCTACAGTGAATG 3643
Db |||||
QY 3205 CGCCAGGCTCAAGCGCGCATGCGCGACGGCGAGGATCTCGTCTGAGCCCATGCGCATGC 3264
Db |||||
QY 3644 GGATGAAGGCGAGCGCGGAGGGGCAAG- -GGCAGATGGCTTTTACAGGGTCTCAG 3701
Db |||||
QY 3265 CTGCTTCGGAATATCATGTGGAAATATGGCGGCTTTTCTGGATTCATCGACTGTGGCG 3324
Db |||||
QY 3702 TACAGCCACCAAGGACACCTACGACGCCCTTTCATGTCAGCGCCCTGCCCCCTCGCTAA- - 3759
Db |||||
QY 3325 GCTGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGA 3384
Db |||||
QY 3760 --CTCGACGGCGCGCGGATCCGGATTTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGT 3817
Db |||||
QY 3385 GCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTC 3444
Db |||||
QY 3818 CCAGGCTCTAGTTTGTACTCAACATATCACCGTGAAGCTAGAGCTATAGATACGAGCCATA 3877
Db |||||
QY 3445 GCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGACGGGACTCTCGGGTTC 3504
Db |||||
QY 3878 GATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAGACCCACCATCTG 3937
Db |||||
QY 3505 GATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGGAATGAAGACCCACCATCTG 3564
Db |||||
QY 3938 TAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCAGAGGCTGG- -AAAATACATAACT 3996
Db |||||
QY 3565 TAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCAGAGGCTGGAAAAATACATAACT 3624
Db |||||
QY 3997 GAGNATAGAGATTTCAAGTCAAGTTAGNACAGA- -GAGNACAGAGATATGGCCAAA 4055
Db |||||
QY 3625 GAGAATAGAGATTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3684
Db |||||
QY 4056 CAGGATATCTGTGTAAGCAGTTCTCTGCCCC- -GCTCAGGGCCAAAGAACAGTTGCAACAGG 4114
Db |||||
QY 3685 CAGGATATCTGTGTAAGCAGTTCTCTGCCCCGCTCAGGGCCAAAGACAGATGGAACAGC 3744
Db |||||
QY 4115 AGAATATGGCCCAAAACAGGATATCTGTGTAAGCAGTTCTCTGCCCCGCTCAGGGCCAAAG 4174
Db |||||
QY 3745 TGAATATGGCCCAAAACAGGATATCTGTGTAAGCAGTTCTCTGCCCCGCTCAGGGCCAAAG 3804
Db |||||
QY 4175 AACAGATGTCCTCCAGATGCGGTCCCGCTCAGCAGTTCTTAGAGAACCATCAGATGTT 4234
Db |||||
QY 3805 AACAGATGTCCTCCAGATGCGGTCCAGCTCAGCAGTTCTTAGAGAACCATCAGATGTT 3864
Db |||||
QY 4235 TCCAGGGTGCCTCCAGAGCCTGAAATGCAACCTGTGCTTATTTGAACTAACCAATCAGTT 4294
Db |||||
QY 3865 TCCAGGGTGCCTCCAGAGCCTGAAATGCAACCTGTGCTTATTTGAACTAACCAATCAGTT 3924
Db |||||
QY 4295 CGTTCTCGCTTCTGTTTCGGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACC 4354
Db |||||
QY 3925 CGTTCTCGCTTCTGTTTCGGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACC 3984
Db |||||
QY 4355 CTTCACTCGCGCGCCAGTCTCTCGATAGACTGCGGTCCCGGGTACCCGTGTTCTCAAT 4414
Db |||||
QY 3985 CTTCACTCGCGCGCCAGTCTCTCGATGACTGAGTTCAGTCCCGGGTACCCGTGTATCCAAT 4044
Db |||||
QY 4415 AAACCTCTTTCAGATGTCATCGACTCGTCTGCTGCTTCTTCTGGAGGGTCT-CTCT 4473
Db |||||
QY 4045 AAACCTCTTTCAGATGTCATCCGACTTGTGCTGCTGCTTCTTCTGGAGGGTCTCTCTCT 4104
Db |||||
QY 4474 GAGTATTGACTACCCCGTCAGCGGGTCTTTTCACTTCTCCCACTACACAGGCTCTCACT 4533
Db |||||


```
Qy 6693 TGCATAATTCCTTACTGTGATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 6752
Db 5933 TGCATAATTCCTTACTGTGATGCGCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACT 5992
Qy 6753 CAACCAAGTCAATCTGAGAAATAGTATGCGCGCACCGAGTTGCTCTTGGCCGCGTCAA 6812
Db 5993 CAACCAAGTCAATCTGAGAAATAGTATGCGCGCACCGAGTTGCTCTTGGCCGCGTCAA 6052
Qy 6813 TACGGGATAATACCGCGCCCATAGCAGAACTTTAAAGTGCTCATCATTTGGAAACGTT 6872
Db 6053 CACGGGATAATACCGCGCCCATAGCAGAACTTTAAAGTGCTCATCATTTGGAAACGTT 6112
Qy 6873 CTTGGGGCGGAAACTCTCAAGATCTTACCGCTGTTGAGATCGAGTTCGATGTAACCCA 6932
Db 6113 CTTGGGGCGGAAACTCTCAAGATCTTACCGCTGTTGAGATCGAGTTCGATGTAACCCA 6172
Qy 6933 CTGCTGACCCAACTGATCTTACGATCTTTTACTTTTACCGAGGTTTCTGGGTGAGCAA 6992
Db 6173 CTGCTGACCCAACTGATCTTACGATCTTTTACTTTTACCGAGGTTTCTGGGTGAGCAA 6232
Qy 6993 AAACAGGAAGCAAAATGCCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAATAC 7052
Db 6233 AAACAGGAAGCAAAATGCCGCAAAAAGGGAATAGGGCGACACGGAATGTTGAATAC 6292
Qy 7053 TCATACCTCTCTTTTCAATATATTATGAAGCATTTATCAGGGTTATGTTCTCATGAGCG 7112
Db 6293 TCATACCTCTCTTTTCAATATATTATGAAGCATTTATCAGGGTTATGTTCTCATGAGCG 6352
Qy 7113 GATACATATTTGAATGATTTAGAAAATAAACAATAAGGGTTCCCGCGCACATTTCCCC 7172
Db 6353 GATACATATTTGAATGATTTAGAAAATAAACAATAAGGGTTCCCGCGCACATTTCCCC 6412
Qy 7173 GAAAAGTGCCACCTGAGCTCTAAGAAACCAATTATTATCATGACATTAACCTATAAAAAATA 7232
Db 6413 GAAAAGTGCCACCTGAGCTCTAAGAAACCAATTATTATCATGACATTAACCTATAAAAAATA 6472
Qy 7233 GGCCTATCAGAGGCCCTTTGCTCT 7257
Db 6473 GGCCTATCAGAGGCCCTTTGCTCT 6497

RESULT 2
AX663075
LOCUS AX663075 9320 bp DNA linear PAT 24-MAR-2003
DEFINITION Sequence 20 from Patent WO02070740.
ACCESSION AX663075
VERSION AX663075.1 GI:29169369
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1. Wiesmueller L.
  Test system for determining gene toxicities
  Patent: WO 02070740-A 20 12-SEP-2002;
  Wiesmueller, Lisa (DE)
  Location/Qualifiers
    1..9320
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="Plasmid p5-Puro-CMV- (N'-EGFP) -CMV-Red- (EGFP-EJ) "
misc_feature
1..1592
/notes="Retroviral Vektor p5NM"
1617..2216
/notes="unnamed protein product; Puromycin-Resistenzgen aus
pRetroOn (Clontech, Palo Alto, CA, USA)"
CDS
1..2216
/transl_table=1
/protein_id="CAD80174.1"
/db_xref="GI:29169370"
/translation="MTBEKPTVRLATRDDVPRAVRTLAAAFADYPATRTVDPDRHIE
```

promoter

CDS

```
RVTELOELFLTRVGLDVGKVVVADGCAAVAVWTTPTSVACAVPAEIGPRMAELSGSR
LAAQQMEGLIAPHRPEKPAWFLATVGVSPDHQKGLGSAVVLDPGVEAAERAGVPAFL
ETAPRNLPFYERLGTFTVADVECPDRATWCMTRKPGA"
2267..2848
/notes="CMV-Promotor aus pEGFP-N1 (Clontech, Palo Alto, CA,
USA)"
2906..3349
/notes="unnamed protein product; N'-EGFP, abgeleitet von
EGFP aus pEGFP-N1 (Clontech, Palo Alto, CA, USA)"
/codon_start=1
/transl_table=1
/protein_id="CAD80175.1"
/db_xref="GI:29169371"
/translation="MWSKGBELFTGVVPIILVELDGDVNGHKFVSVSGEGEDATYVKLT
LKPICTTGKLPVPMPTLVTLTYGVQCFSPYDHMKQHDFFKSGAMPEGYVQERTIFFK
DDGNYKTRAEVKEFGEDTLVNRIELKGIDFKEDGNILGHKLEYNVR"
3374..3392
/notes="Retroviral Vektor p5NM"
3411..3992
/notes="CMV-Promotor aus pDeRedi-N1 (Clontech, Palo Alto,
CA, USA)"
4038..4718
/notes="unnamed protein product; Red aus pDeRedi-N1
(Clontech, Palo Alto, CA, USA)"
/codon_start=1
/transl_table=1
/protein_id="CAD80176.1"
/db_xref="GI:29169372"
/translation="MVRSSKNVIKEPMRPFVMEGTVNGHEFIEGEGEGRPYRGHNT
VKLVYKGGPLPANDILSPQFOYGVSKVYVHPADIPDYKLSLFPPEGFKWVNFED
GVYTVQDSLDQGCIFYKFIQVNPDSGPMVKTKTGWNEASTBRLIYPRDGVKKG
EIHKALKDKDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYITVEQYRT
EGRHILFL"
4766..5509
/notes="EGFP-EJ, abgeleitet von EGFP aus pEGFP-N1
(Clontech, Palo Alto, CA, USA)"
5527..9320
/notes="Retroviral Vektor p5NM"
```

misc_feature

misc_feature

ORIGIN

```
Query Match 43.8%; Score 3349; DB 6; Length 9320;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 3570; Conservative 0; Mismatches 150; Indels 61; Gaps 7;

Qy 3772 GCGGATCCGGATAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTT 3831
Db 5541 GGGCCCGCGATAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTT 5600
Qy 3832 TGACTCAACAATATCACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAGA 3891
Db 5601 TGACTCAACAATATCACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAGA 5660
Qy 3892 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCACCCTGTAGTTTGGCAAGC 3951
Db 5661 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCACCCTGTAGTTTGGCAAGC 5720
Qy 3952 TAGCTTAAGTAAGCCCATTTTTCGAAGGCATGGAATAATACATACTGAGAATAGAGAAGTT 4011
Db 5721 TAGCTTAAGTAAGCCCATTTTTCGAAGGCATGGAATAATACATACTGAGAATAGAGAAGTT 5780
Qy 4012 CAGATCAAGTTTAGGAACACAGAGACAGAGATATATGGGCCAAACAGAGATATCTGTGTA 4071
Db 5781 CAGATCAAGTTTAGGAACACAGAGACAGAGATATATGGGCCAAACAGAGATATCTGTGTA 5840
Qy 4072 AGCAGTTCCTGCCCCC-GCTCAGGGCCAAAGAACAGTTGGACAGAGATATGGGCCAAAC 4130
Db 5841 AGCAGTTCCTGCCCCCCTCAGGGCCAAAGAACAGTTGGACAGAGATATGGGCCAAAC 5900
Qy 4131 AGGATATCTGTGTAGAGCAGTTCTCTGCCCGGGCTCAGGGCCAAAGAACAGATGGTCCCAAG 4190
Db 5901 AGGATATCTGTGTAGAGCAGTTCTCTGCCCGGGCTCAGGGCCAAAGAACAGATGGTCCCAAG 5960
Qy 4191 ATCGGGTCCCGCCCTCAGCAGTTTCTTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAG 4250
```


Qy 6353 GCCTCCAGATTATCAGCAATTAACACGACGCGGAGAGGGCCGAGCGGAGAGTGGTCC 6412
Db 8180 GCCTCCAGATTATCAGCAATTAACACGACGCGGAGAGGGCCGAGCGGAGAGTGGTCC 8239
Qy 6413 TGCAACTTTATCCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGGCTAGAGTAAGTAG 6472
Db 8240 TGCAACTTTATCCCGCTCCATCCAGTCTATTAAATTTGTCGGGAGGCTAGAGTAAGTAG 8299
Qy 6473 TTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCAATTCGTACAGGC-TCGTGGTGTACAG 6531
Db 8300 TTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCAATTCGTACAGGCATCGTGGTGTACAG 8359
Qy 6532 CTCGTGTTTGGTATGCGTTCATTCAGCTCCGGTTCCTCAAGCATCAAGCGAGTTACATG 6591
Db 8360 CTCGTGTTTGGTATGCGTTCATTCAGCTCCGGTTCCTCAAGCATCAAGCGAGTTACATG 8419
Qy 6592 ATCCCCCATGTTGTCACAAAGCGTTAGCTCTTCCTCGGTCTCGATCGTTCAGAG 6651
Db 8420 ATCCCCCATGTTGTCACAAAGCGTTAGCTCTTCCTCGGTCTCGATCGTTCAGAG 8479
Qy 6652 TAAGTTGCCCGCAGTGTATCACTCATGTTATGCGCAGCTGCAATAATCTCTTACTGT 6711
Db 8480 TAAGTTGCCCGCAGTGTATCACTCATGTTATGCGCAGCTGCAATAATCTCTTACTGT 8539
Qy 6712 CATGCCATCCGTAAGAGCTTTTCTGCACTGGTGAGTACTCAACCAAGTCAATCTGAGA 6771
Db 8540 CATGCCATCCGTAAGAGCTTTTCTGCACTGGTGAGTACTCAACCAAGTCAATCTGAGA 8599
Qy 6772 ATAGTGTATGCGGACGAGTGTCTTCTCCCGCGCTCAATACGGGATATACCGGCC 6831
Db 8600 ATAGTGTATGCGGACGAGTGTCTTCTCCCGCGCTCAATACGGGATATACCGGCC 8659
Qy 6832 ACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAGCTTCTCGGGCGCAAACTCTC 6891
Db 8660 ACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAAGCTTCTCGGGCGCAAACTCTC 8719
Qy 6892 AAGGATCTTACCGTGTGAGATCAGTTCGATGTAAACCACTCGTGACCACTGATC 6951
Db 8720 AAGGATCTTACCGTGTGAGATCAGTTCGATGTAAACCACTCGTGACCACTGATC 8779
Qy 6952 TTCAGCATCTTTTACTTTTACAGGCTTTCAGGCTTTCAGGCTGAGCAAAAGGCAAAATGC 7011
Db 8780 TTCAGCATCTTTTACTTTTACAGGCTTTCAGGCTGAGCAAAAGGCAAAATGC 8839
Qy 7012 CGCAAAAGGGAATAAGGCGGACACGAAATGTTGAATATCTACTCTTCTTTTCA 7071
Db 8840 CGCAAAAGGGAATAAGGCGGACACGAAATGTTGAATATCTACTCTTCTTTTCA 8899
Qy 7072 ATATTATTGAAGCATTTATCAGGTTATTGTTCTCATGACGCGATACATATTGAAATGAT 7131
Db 8900 ATATTATTGAAGCATTTATCAGGTTATTGTTCTCATGACGCGATACATATTGAAATGAT 8959
Qy 7132 TTAGAAAATAAACAATAGGGTTCGCGGACATTTCCCGGAAAGTGCCACCTGACGT 7191
Db 8960 TTAGAAAATAAACAATAGGGTTCGCGGACATTTCCCGGAAAGTGCCACCTGACGT 9019
Qy 7192 CTAAGAACCAATTTATCATGACATTAACCTATATAAATAGGCGTATCAGAGGCCCTT 7251
Db 9020 CTAAGAACCAATTTATCATGACATTAACCTATATAAATAGGCGTATCAGAGGCCCTT 9079
Qy 7252 TCGTCTCGCGGCTTTCGGTGTATGACGCTGAAAACTCTGACACATGACGCTCCCGGAGAC 7311
Db 9080 TCGTCTCGCGGCTTTCGGTGTATGACGCTGAAAACTCTGACACATGACGCTCCCGGAGAC 9139
Qy 7312 GGTCAAGCTTGTCTGTAAGCGATGCGGAGCAGACAAAGCCGTGAGGCGCGTCAAGC 7371
Db 9140 GGTCAAGCTTGTCTGTAAGCGATGCGGAGCAGACAAAGCCGTGAGGCGCGTCAAGC 9199
Qy 7372 GGGTGTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAG 7431
Db 9200 GGGTGTGGCGGGTGTGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGACTGAG 9259

Qy 7432 AGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAG 7491
Db 9260 AGTGACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAG 9319
Qy 7492 G 7492
Db 9320 G 9320
RESULT 3
LOCUS E23356 6644 bp DNA linear PAT 18-JUN-2001
DEFINITION Virus vector system expressing apoptosis-related gene.
ACCESSION E23356
VERSION E23356.1 GI:13024379
KEYWORDS JP 199075859-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 6644)
AUTHORS Hirofumi, H.
TITLE Virus vector system expressing apoptosis-related gene
JOURNAL Patent: JP 199075859-A 2 23-MAR-1999;
R B R JENSERU KK
COMMENT OS Unidentified
PN JP 199075859-A/2
PD 23-MAR-1999
PF 08-SEP-1997 JP 1997259235
PR HIROFUMI HAMADA
PI C12N15/09, C12N5/10, C12N7/00//A61K35/76, A61K48/00, (C12N5/10, PC
C12R1:91),
PC (C12N7/00, C12R1:92), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT 1..6644 /organism='Unidentified'.
FEATURES
source
1..6644 /organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'
ORIGIN
Query Match 43.2%; Score 3307.8; DB 6; Length 6644;
Best Local Similarity 72.9%; Pred. No. 0;
Matches 4739; Conservative 0; Mismatches 1352; Indels 410; Gaps 21;
Qy 843 TTCTAGAGAACCATCAGATGTTCCAGGTGCCCCAAGGACCTGAAATGACCTGTGCT 902
Db 467 TGCTAGAGAACCATCAGATGTTCCAGGTGCCCCAAGGACCTGAAATGACCTGTGCT 526
Qy 903 TATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTTGCGCGCTTCTGCTCCCCGAG 962
Db 527 TATTGAACTAACCAATCAGTTCGCTTCTGCTTCTGTTGCGCGCTTCTGCTCCCCGAG 586
Qy 963 CTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGGACCTCCCGATAGACTGCGTCCG 1022
Db 587 CTCAATAAAGAGCCCAACCCCTCACTCGGCGGCGGACCTCCCGATAGACTGCGTCCG 646
Qy 1023 CCCGGGTACCGGTATTCCCAATAAGCTCTTGTGTTTGCATCCGAATCGTGACTCGC 1082
Db 647 CCCGGGTACCGGTATTCCCAATAAGCTCTTGTGAG-ITGCAATCGACTTGTGCTCGC 705
Qy 1083 TGATCTTGGAGGGTCTCTCAGATTGATGACTGCC-ACCTCGGGGCTTTTCA TTT 1141
Db 706 TGTTCTTGGAGGGTCTCTCTCTGAGTGATTGACTACCCGTCAGCGGGGCTTTTCA TTT 765
Qy 1142 GGAGTTCCACCGAGATTGGAGACCCCTCGCCAGGACCAACCCACCCCGCGGGAG 1201
Db 766 GGGGGCTCGTCCGGGATCGGAGACCCCTCGCCAGGAGCAACCGA-CCCACACCGGGAG 824

Db 2689 CTAGGCCCCCGAA-----CCACGGGACGTGGTTTCTCTTTGAAAAACACGATATA 2741
Qy 3421 AGAGTGAAGTTTCAGAGAGCGCAGAGCCCTCCCGCTACACGAGCGGCCAGAACAGCTC 3480
Db 2742 CCATGGTTCATGAAACATTTAAACATTTCTCAACAGATCTAGAATTAGTAGAAGTAGCGA 2801
Qy 3481 TATAACGAGCTCAATCTAGGACGAGAGAGAGAGTACGATGTTTGGACAGAGACGTGGC 3540
Db 2802 CAGAGAAGATTACAAATGCTTTTATGAGGATAATAACATCATGTGGGAGCGCAATTGCTA 2861
Qy 3541 CGGACCTCTGAGATGGGGGAAAGCCGAGAGAGGAAACCCCTCAGGAAGGCTCTGACAA 3600
Db 2862 CGAATAACAGGAGAAATCATTTCCGACGTACATATTTGAAGCGTATATAGGACGATAGCTG 2921
Qy 3601 GAATCGAGAAAGATAGATGGCGGAGCCCTACAGTGAGATGTTGGATGAAAGCGAGCGC 3660
Db 2922 TTTGTGCAGAGCCATTCGATTTGGTAG--TGCAGTTTTCGAATGCAACAAAGGATTTTGA 2979
Qy 3661 CGGAGGGCAGGGCCAGATGGCTTTTACCAGGCTCTCAGTACAGCCACCAGGACACC 3720
Db 2980 CACGATTGTAGCTGTAGACACCCCTTATTTCTGACGAAGTAGATAGAAGTATTCAGTGGT 3039
Qy 3721 TACGACGCCCTTCACATGCGAGCCCTGCCCTCGCTAACTCGACGCGCGCGGATCCG 3780
Db 3040 AAGTCCTTGTGTATGTAGG-----GAGTTGATTTAGACTATGACCA 3085
Qy 3781 GATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTAAC 3840
Db 3086 GATTGTTTGTGTTAATAGAAATGAATGGCAAGTTAGTCAAACTACGATTGAAGAACTC 3145
Qy 3841 AATATCACCAGCTGAGCCTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900
Db 3146 ATTCCACTCAATATATACC--GAAATTAAGTGCAGCATATAAATAAAGATTTTATTTA 3203
Qy 3901 GTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAG 3960
Db 3204 GTCTCCAGAAAAAGGGGGAATGAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAG 3263
Qy 3961 TAACGCCATTTTCAAGGCATGG--AAAATACATAACTGAGAAATAGAGAAGTTTCAGATCAA 4019
Db 3264 TAACGCCATTTTCAAGGCATGGAAAAATAATAACTGAGAAATAGAGAAGTTTCAGATCAA 3323
Qy 4020 GGTTAGGAACAGA--GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 4078
Db 3324 GGTGAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 3383
Qy 4079 CCTGCCCC--GCTCAGGGCCAAAGACAGTTGGAAACAGAGAAATATGGGCCAAACAGGATAT 4137
Db 3384 CCTGCCCCGCTCAGGGCCAAAGACAGTGAATATGGGCCAAACAGGATAT 3443
Qy 4138 CTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGCGGT 4197
Db 3444 CTGTGTAAGCAGTTCTGCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGCGGT 3503
Qy 4198 CCGGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCAGGGTGGCCCCAGACCTGA 4257
Db 3504 CCAGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCAGGGTGGCCCCAAGACCTGA 3563
Qy 4258 AATGACCTGTCCTTATTTGAACTTAACCAATCAGTTGCTTCTGCTTCTGTTTCGGCG 4317
Db 3564 AATGACCTGTCCTTATTTGAACTTAACCAATCAGTTGCTTCTGCTTCTGTTTCGGCG 3623
Qy 4318 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCCTCAGCGCGGCCAGTCTCTC 4377
Db 3624 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCCTCAGCGCGGCCAGTCTCTC 3683
Qy 4378 CGATAGACTGGTTCGCGCGGTAACCGTGTCTCAATAAACCTCTTTCGATTTGCATCCG 4437
Db 3684 CGATTGACTGAGTCGCGCGGTACCCGTGTATCCAATAAACCTCTTTCGATTTGCAAAA 3743
Qy 4438 ACTCGTGTCTCGTGTCTTCTGGGAGGGTCTCTCTGAGTATGACTTACCCGTCAGGG 4497
Db 3744 AA 3803

Qy 4498 GGTCTTTTCAGTTCTCCCACTACACAGGCTCTCACTAAACATTCCTGATGTCCCGCAGGGA 4557
Db 3804 AA 3863
Qy 4558 CTCGTCAGCCCGTGTGTTTATTAATAAATCAAGACAGTGTTCCTTCAAGCCAG 4617
Db 3864 AA 3923
Qy 4618 ACTACATCTCTGACTCTCGGCTTTATAAAGAATGTTGAAGGGCTCTGTGGACTATCTGCC 4677
Db 3924 AA 3983
Qy 4678 ACAGACTTTTAAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATCTCTCGT 4737
Db 3984 AA 4043
Qy 4738 CTATTTTGTGCTGCTCTCCGTTTAAATTTCTAGTTTGCACCTCTCTCTGAGAGCAC 4797
Db 4044 AA 4103
Qy 4798 GGCATTGCGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTTGCCTGGCT 4857
Db 4104 AA 4163
Qy 4858 CAGTGTGAGATTTTGCATATAAAGGGGTCTGCTGCCCTGTGTACAGACAGATCGGAATCT 4917
Db 4164 AA 4223
Qy 4918 AGAGTGCATACTCAGAGTCCCCGGGTTCCGGGCTCTGATCTCAGGCGATCTTTGCCCTA 4977
Db 4224 AA 4283
Qy 4978 GAGATCTCTACGCGCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCT 5037
Db 4284 AA 4343
Qy 5038 CATAGCTGTTTCTGTGTGAATTTGTTATCGCTCACAATTTCCACACATATAGAGCG 5097
Db 4344 AA 4403
Qy 5098 GAGCATTAAGTGTAA-----AGCCT 5118
Db 4404 AA 4463
Qy 5119 GGGTGCCTAATGAGTCACTCACTCAATTAATTTGGTTTGGCTCACTGCTCC----- 5171
Db 4464 AGNAGCTTAAATGAGAAAAAATCACTGGATATACCACTGATATATCCCAATGGC 4523
Qy 5172 -----GCTTTCCAGTCGGGAAACCTGCTGT 5196
Db 4524 ATCGTAAAGAACATTTTGAAGGCATTTTCAGTCAGTTGCTCAATGTACTATAACACAGCCG 4583
Qy 5197 GCCAGCTGCAATTAATGATCGGCCAACCGCGGGGAGAGGGGTTTCGTTATTCGGCGCT 5256
Db 4584 TTCACTGCTCAATTAATGATCGGCCAACCGCGGGGAGAGGGGTTTCGTTATTCGGCGCT 4643
Qy 5257 CTTCCGCTTCTCCTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5316
Db 4644 CTTCCGCTTCTCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4703
Qy 5317 CAGTCTCACTCAAGGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACACAGGAAGA 5376
Db 4704 CAGTCTCACTCAAGGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACACAGGAAGA 4763
Qy 5377 ACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGCTTGTGCGGT 5436
Db 4764 ACATGTAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGCTTGTGCGGT 4823
Qy 5437 TTTTTCATAGCTCCGCCCTCGACGAGCATCAAAAAATCAGAGCTCAAGTCAGAGGT 5496
Db 4824 TTTTTCATAGCTCCGCCCTCGACGAGCATCAAAAAATCAGAGCTCAAGTCAGAGGT 4883

QY	5497	GGCGAAACCCGACGACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGC	5556	Db	5964	CAAGCGAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTAGCTCTTCGGTCTCTC	6023
Db	4884	GGCGAAACCCGACGACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGC	4943	QY	6636	CGATCGTTGTCAAGTAAGTTGGCCGACAGTGTATCACTCATGCTTATGGCAGCACTGC	6695
QY	5557	GCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCCGGAA	5616	Db	6024	CGATCGTTGTCAAGTAAGTTGGCCGACAGTGTATCACTCATGCTTATGGCAGCACTGC	6083
Db	4944	GCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCCGGAA	5003	QY	6696	ATAATTTCTTACTCTCATGTCATGCGCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAA	6755
QY	5617	GGTGGCGCTTTCTCATAGCTCAAGCTGAGTGTAGTACTCACTAGTTGCGGTGAGTGTGCGT	5676	Db	6084	ATRAATTTCTTACTCTCATGTCATGCGCATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAA	6143
Db	5004	GGTGGCGCTTTCTCAATGCTCACGCTGAGTGTAGTACTCACTAGTTGCGGTGAGTGTGCGT	5063	QY	6756	CCAAGTCATTTCTGAGAATAGTGTATGCGCGCAACCTTTAAAGTGTCTCATCATTTGGAACACGTTCTT	6815
QY	5677	CCAAGTCGCGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTGCGCTTTATCCGTA	5736	Db	6144	CCAAGTCATTTCTGAGAATAGTGTATGCGCGCAACCTTTAAAGTGTCTCATCATTTGGAACACGTTCTT	6203
Db	5064	CCAAGTCGCGCTGTGTGACGAACCCCGCTTTCAGCCCGACCGCTGCGCTTTATCCGTA	5123	QY	6816	GGGATAATACCGCGCCACATAGCAAACTTTAAAGTGTCTCATCATTTGGAACACGTTCTT	6875
QY	5737	ACTATGCTTGAAGTCAACCCCGGTAAGACACGACTTATCGCCACTGGCAGACGCACTG	5796	Db	6204	GGGATAATACCGCGCCACATAGCAAACTTTAAAGTGTCTCATCATTTGGAACACGTTCTT	6263
Db	5124	ACTATGCTTGAAGTCAACCCCGGTAAGACACGACTTATCGCCACTGGCAGACGCACTG	5183	QY	6876	CGGGCGGAAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC	6935
QY	5797	GTAACAGGATTACGAGCGAGTATGTAGGCGGTGCTACAGATGTTCTGAAGTGTGCG	5856	Db	6264	CGGGCGGAAAACCTCTCAAGGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC	6323
Db	5184	GTAACAGGATTACGAGCGAGTATGTAGGCGGTGCTACAGATGTTCTGAAGTGTGCG	5243	QY	6936	GTGCAACCAACTGATCTTTTACGATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAA	6995
QY	5857	CTAACTACGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA	5916	Db	6324	GTGCAACCAACTGATCTTTTACGATCTTTTACTTTTCCAGCGCTTTCTGGGTGAGCAAAA	6383
Db	5244	CTAACTACGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTA	5303	QY	6996	CHGGAAGGCAAAATCCGCAAAAAGGAATAAGGCGCACACGGAATAATGTTGAATACTCA	7055
QY	5917	CCTTCGGAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCAGCTGCTAGCGGTG	5976	Db	6384	CHGGAAGGCAAAATCCGCAAAAAGGAATAAGGCGCACACGGAATAATGTTGAATACTCA	6443
Db	5304	CCTTCGGAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACCCAGCTGCTAGCGGTG	5363	QY	7056	TACTCTCTCTTTTCAATATTAATCAAGCATTTATCAGGTTATTTGTTCTCATGAGCGAT	7115
QY	5977	GTTTTTTTGTTCGACGACGAGATTACGCGCAGAAAAAGATCTCAAGAGATCTCTT	6036	Db	6444	TACTCTCTCTTTTCAATATTAATCAAGCATTTATCAGGTTATTTGTTCTCATGAGCGAT	6503
Db	5364	GTTTTTTTGTTCGACGACGAGATTACGCGCAGAAAAAGATCTCAAGAGATCTCTT	5423	QY	7116	ACATATTTGAATGTTTGTAGAAAAATAACAAATAGGCGTTCGCGCACATTTTCCCGGAA	7175
QY	6037	TGATCTTTTCTACGGGCTCTGACGCTCAGTGGAACGAAACCTCACGTTAAGGGATTTGG	6096	Db	6504	ACATATTTGAATGTTTGTAGAAAAATAACAAATAGGCGTTCGCGCACATTTTCCCGGAA	6563
Db	5424	TGATCTTTTCTACGGGCTCTGACGCTCAGTGGAACGAAACCTCACGTTAAGGGATTTGG	5483	QY	7176	AGTGCCACCTGACGCTCTAAGAACCATTTATCATGACATTAACCTTAATAATAGGC	7235
QY	6097	TCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAATTTTA	6156	Db	6564	AGTGCCACCTGACGCTCTAAGAACCATTTATCATGACATTAACCTTAATAATAGGC	6623
Db	5484	TCATGAGATTATCAAAAAGGATCTTCACTAGATCTTTTAAATTAATAAATGAATTTTA	5543	QY	7236	GTATCAGGAGGCCCTTTGCTC 7256	
QY	6157	ATCAATCTAAAGTATATGATGATAAATTTGCTGACAGTTACCAATGCTTAATCAGTG	6216	Db	6624	GTATCAGGAGGCCCTTTGCTC 6644	
Db	5544	ATCAATCTAAAGTATATGATGATAAATTTGCTGACAGTTACCAATGCTTAATCAGTG	5603	RESULT 4			
QY	6217	AGGCACCTATCTCAGCGATCTGTCTATTTTCTATTTCTATCCATAGTTGCTGACTCCCGCTGC	6276	AX823827			
Db	5604	AGGCACCTATCTCAGCGATCTGTCTATTTTCTATTTCTATCCATAGTTGCTGACTCCCGCTGC	5663	LOCUS	AX823827	6505 bp DNA linear PAT 12-DEC-2003	
QY	6277	TGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGC	6336	DEFINITION	Sequence 5 from Patent WO03070958.		
Db	5664	TGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGC	5723	ACCESSION	AX823827		
QY	6337	GAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAACACGACCGCGAAGGCGC	6396	VERSION	AX823827.1	GI:39750146	
Db	5724	GAGACCCAGCTCACCGCTCCAGATTTATCAGCAATAACACGACCGCGAAGGCGC	5783	KEYWORDS	murine stem cell retroviral vector (MSCV)		
QY	6397	AGCGAGAAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGG	6456	SOURCE	murine stem cell retroviral vector		
Db	5784	AGCGAGAAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTATTTGTCGCGG	5843	ORGANISM	murine stem cell retroviral vector		
QY	6457	ARAGTAGATAAGTGTGCGCAAGTAAATGATTGCGCAACGTTGTGCAATGCTTACAG	6516	REFERENCE	other sequences; artificial sequences; vectors.		
Db	5844	ARAGTAGATAAGTGTGCGCAAGTAAATGATTGCGCAACGTTGTGCAATGCTTACAG	5903	AUTHORS	1		
QY	6517	GC-TGCTGTGTACGCTCGTGTGTTGTTGATGGCTTCAATCAGCTCCGCTTCCCAACGAT	6575	TITLE	Yao, S. and Ellis, J.		
Db	5904	GCATCGTGTGTACGCTCGTGTGTTGTTGATGGCTTCAATCAGCTCCGCTTCCCAACGAT	5963	JOURNAL	Retroviral gene therapy vectors including insulator elements to provide high levels of gene expression		
QY	6576	CAAGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCTC	6635	FEATURES	Patent: WO 03070958-A 5 28-AUG-2003; THE HOSPITAL FOR SICK CHILDREN (CA)		
				source	Location/Qualifiers		
					1. .6505		
					/organism="murine stem cell retroviral vector"		
					/mol_type="unassigned DNA"		
					/db_xref="taxon:258023"		
					1. .6505		
					/note="n = g, a, t, or c."		
				LTR			
				ORIGIN			

Query Match	42.4%;	Score 3244.4;	DB 6;	Length 6505;
Best Local Similarity	71.1%;	Prod. No. 0;		
Matches 5064;	Conservative	0;	Mismatches 1020;	Indels 1035; Gaps 25;
Qy	552	TGAAGACCCACCTGTAGGTTTCGCAAGCTAGCTTAAGTAACGCCATTTTTCGAAGCAT	611	
Db	1	TGAAGACCCACCTGTAGGTTTCGCAAGCTAGCTTAAGTAACGCCATTTTTCGAAGCAT	60	
Qy	612	GGAAATAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTTTAGGAAACAGAGACACGCA	671	
Db	61	GGAAATAATACATACTGAGAAATAGAGAAGTTTCAGATCAAGGTTTAGGAAACAGAGACACGCA	119	
Qy	672	GAATATGGGCCAAACAGAGATATCTGTGTAAGACAGTTCTCTGCCCGCTCAGGGCCAGAA	731	
Db	120	-----	119	
Qy	732	CAGTTTGGAAACAGGAGATATATGGGCCAAACAGAGATATCTGTGTAAGACAGTTCTCTGCCCGCT	791	
Db	120	-----AGAAATATGGGCCAAACAGAGATATCTGTGTAAGACAGTTCTCTGCCCGCT	166	
Qy	792	GCTCAGGGCCAAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCAGGTTTCTAGAGA	851	
Db	167	GCTCAGGGCCAAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCAGGTTTCTAGAGA	226	
Qy	852	ACCATCAGATGTTTTCAGGGTGCCCCAAGACCTG-AAATGACCTGTGCTTATTTGAA	910	
Db	227	ACCATCAGATGTTTTCAGGGTGCCCCAAGACCTGAAATGACCTGTGCTTATTTGAA	286	
Qy	911	CTAACCAATCAGTTTCGCTTCTGTTTTCGCGCGCTTCTGCTCCCGAGCTCAATAA	970	
Db	287	CTAACCAATCAGTTTCGCTTCTGTTTTCGCGCGCTTCTGCTCCCGAGCTCAATAA	346	
Qy	971	AAGAGCCCAACAAACCTCAGTTCGGCGGCCAGTCTCTCGATAGACTGCGTCCCGCGGTA	1030	
Db	347	AAGAGCCCAACAAACCTCAGTTCGGCGGCCAGTCTCTCGATAGACTGCGTCCCGCGGTA	406	
Qy	1031	CCGCTATTTCCAAATAAGCTCTTGTGTTTTCGATTCGGAATCGTGAAGTCTGCTGATCTT	1090	
Db	407	CCGCTATTTCCAAATAAGCTCTTGTGTTTTCGATTCGGAATCGTGAAGTCTGCTGATCTT	466	
Qy	1091	GGGAGGTCCTCAGATTGATTCAGTCCCACTCGGGGCTCTTCAATTTGGAGGTTCC	1150	
Db	467	GGGAGGTCCTCAGATTGATTCAGTCCCACTCGGGGCTCTTCAATTTGGAGGTTCC	526	
Qy	1151	ACCGAGATTTGGAGACCCCTGCCAGGAGACACCGACCCCGCCCGGAGGTAAGCTGG	1210	
Db	527	ACCGAGATTTGGAGACCCCTGCCAGGAGACACCGACCCCGCCCGGAGGTAAGCTGG	586	
Qy	1211	CCAGCAATATCTGTGTCGTGTC-----CGATTGTAGTGTCTATGATCATTTTA	1262	
Db	587	CCAGCGTCGTTCGTGTCGTCTCTGTCTTGTGCGGTGTTGTGCGGCACTAATGTT	646	
Qy	1263	TGCGCTCGCTCGGTACTAGTTAGTACTACTAGCTCTGTATCTGCGGACCCGTGGTGGAA	1322	
Db	647	TGCGCTCGCTCGGTACTAGTTAGTACTACTAGCTCTGTATCTGCGGACCCGTGGTGGAA	706	
Qy	1323	CTGACGAGTTTCGGAACACCCCGCCGCAACCTCGGAGACGTCCACAGGACTTCGGGGGCC	1382	
Db	707	CTGACGAGTTTCGGAACACCCCGCCGCAACCTCGGAGACGTCCACAGGACTTCGGGGGCC	766	
Qy	1383	GTPTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTGTAGACTCTTTGTGTGCAACC	1442	
Db	767	GTPTTTGTGGCCGACCTGAGGAGGGAGTCTGATGTGGAATCCGACCC-----	814	
Qy	1443	CCCTTAGAGGAGGATATGTGTTCTCGTAGGAGACGAGAACCTTAAACAGTTTCCCGCT	1502	
Db	815	-----CGTCAGGATATGTGTTCTCGTAGGAGACGAGAACCTTAAACAGTTTCCCGCT	867	
Qy	1503	CCGTCTGAAATTTTGTCTTTTCGTTTGTGGAACGGAAGCCGCGCGCTGTGT---CTGC	1559	
Db	868	CCGTCTGAAATTTTGTCTTTTCGTTTGTGGAACGGAAGCCGCGCGCTGTGTCTGCTGACGGC	927	
Qy	1560	TGCAGCATCGTTCGTGTTGTCTGCTGTGATGCTGTGTTCTGTATTTGTCTGAAAAATATG	1619	

928	Db	 TGCAGCATCGTTCGTGTGTCTCTGTCTGACTGTGTCTCTGTATTTGTCTGAAATTAG	987
1620	Qy	GGCCCGGGCTAGACTGTTTACACATCCCTTAAGTTTGACCTTAGTCACTCGGAAGATGC	1679
988	Db	GGCC-----AGACTGTTACCACTCCCTTAAGTTTGACCTTAGTCACTCGGAAGATGC	1041
1680	Qy	GAGCGGATCGCTCAACAACCACTCGGTAGATGTCACAAAGAGAGAGTGGGTACCTTCTGC	1739
1042	Db	GAGCGGATCGCTCAACAACCACTCGGTAGATGTCACAAAGAGAGAGTGGGTACCTTCTGC	1101
1740	Qy	TCTCGAAGATGGCCAACTTTAAGTCGGATGGCGCGGAGACGGCACCTTTAAACCGAGAC	1799
1102	Db	TCTCGAAGATGGCCAACTTTAAGTCGGATGGCGCGGAGACGGCACCTTTAAACCGAGAC	1161
1800	Qy	CTCATCAACCAGGTTAAGATCAAGGTCCTTTTACCTTGGCCCGCATGGACACCCAGACCAG	1859
1162	Db	CTCATCAACCAGGTTAAGATCAAGGTCCTTTTACCTTGGCCCGCATGGACACCCAGACCAG	1221
1860	Qy	GTCCCTTACATCTGTGACCTCGGAAGCCTTGGCTTTTGTACCCCTCTCCTGGGTCAAAGCCC	1919
1222	Db	GTCCCTTACATCTGTGACCTCGGAAGCCTTGGCTTTTGTACCCCTCTCCTGGGTCAAAGCCC	1281
1920	Qy	TTTGTACACCTTAAGCCTCGGCTCTCTTCTTCCTCAATCGCGCCCGTCTCTCCCTTTGAA	1979
1282	Db	TTTGTACACCTTAAGCCTCGGCTCTCTTCTTCCTCAATCGCGCCCGTCTCTCCCTTTGAA	1341
1980	Qy	CCTCCTCGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGC	2039
1342	Db	CCTCCTCGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTAGGC	1401
2040	Qy	GCCCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCT	2099
1402	Db	GCCGGAT-----TCGTTAACCTCGAG	1423
2100	Qy	GACCTTGACATGACAAGAGTTACTAACAGCCCTCTCTCCAAAGTCACTTTACAGGCTTCT	2159
1424	Db	AGATCTAATTCTACCGGTTAGGGAGGCGCTTTTCCAAAGCGAGTCTGGAGCATGGCGTT	1483
2160	Qy	ACTTAGTCCAGCAGAGTCTGGAGACCTCTGGCGGACGCTTACCAGAACAACTGGACC	2219
1484	Db	TAGCAGCCCGCTGGCACTTGGCGCTTACAACAATGGCGCTCTGGCGCTCGCACATTTCCAC	1543
2220	Qy	GACCGGTGGTACCTTCACTTACCGAGTCCGGCACACAGTGTGGGTCCGCCGACACAGA	2279
1544	Db	ATCCACCGGTAGCGC-----AACCGGCTCCGTTCTTTGGTGGCCCTTCGCGCCA	1594
2280	Qy	CTAAGAACCTTAGAACCTTCGTGTGAAGAGACCTTTACAGCTCTGTCTGTGAACACCCCAACG	2339
1595	Db	CTTCTACTCC---TCCCTTAGTCAGAAAGTTTCCCCAGCAAGCTCGGCTCGTCGACGGA	1650
2340	Qy	CCCTCAAGTAGACGGCATCGCAGCTTGGATACACGCGCCCACTGAGAGGCTCGCGACC	2399
1651	Db	CGTGACAAATGGGAAGTAGACGCTTCACTAGTCTCTGTGCAGATGGACAGACCGCTGAGC	1710
2400	Qy	CCGGGGGTGACCATTCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTTGGT	2459
1711	Db	AAATGAAGCGGTAGGCCCTTTGGGCGACGGCCCAATAGCAGCTTTGTTCTTCGCTTTCT	1770
2460	Qy	AGCAACAGCTTACAGGTGTCCACTTCGACATCTCAGCTGACCCAGAGCCCAAGCAAGCCTGAG	2519
1771	Db	GGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGCGGGCTCAG-----	1813
2520	Qy	CGCCAGCGTGGGTGACAGAGTGACCATCATCTGTAGGGCCAGTTCAGATGTGGGTACTTC	2579
1814	Db	-----GGCGGGGTTCAGGGCGGGCGGGCGGCCGAAGGTCCTCCGAGGCCCGGCATCT	1868
2580	Qy	TGTAGCTTGTGATCAGCAGAAAGCAGTAAGGCTCCAAAGCTCTCATCTGATCTACTTGGACATC	2639
1869	Db	GCAGCTTCAAAGCGCATGCTCTGCCGCGCTGTTCTCTTCTCTCTCATCTCCGGGCTTT	1928
2640	Qy	CACCCGCACTGTGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGATTTCCAC	2699

Db 1929 CGACCTGCAGCCAAATATATGGGATCGGCATTTGAACAAGATGGATGTGACGACAGTTCTC-- 1986
Qy 2700 CTTTACCATCAGACGCTCCAGCCAGAGACATCGCCACTACTACTGCGCAGCAATATAG 2759
Db 1987 -----CGGCCGCTTGGGTGGAGAGGCTAT--TCGGCTATGACTGGGCAACAAGAC 2035
Qy 2760 CCTTATCGGTGTTCGGCCAGGAGGACCAAGGTGGAAATCAAAAGAGGTGGCTCAGGATC 2819
Db 2036 ATTCGGCTGCTC-----TGATGCCCGGTGT 2062
Qy 2820 GGGTGGATCCGCTCTGTGGGTCTCAGGATCGAGGTCCAACTGGTGGAGAGCGGTGGAGG 2879
Db 2063 CCGGCTGTTCAGCGCAGGGGCGCCGGTCTTTTGTTCAGAACCGACCTGTCCGGTGCCT 2122
Qy 2880 TGTGTGCAACTGGCCGCTCCCTGGCTGTCTGCTCCGATCTGGCTTCGATTTAC 2939
Db 2123 GAATGAACATGCAGGACGAGCGCGCTATCGTGGCTGGCCACGACGGGCTTCCCTTG 2182
Qy 2940 CACATATTGGATGAGTTGGGTGAGACAGGCACCTGGAAAGGTCTTGGATGGATTTGGAGA 2999
Db 2183 CGCAGCTGTCTCGACGTGTCTACTGAAGCGGAGGAGTGGCTGCTATTGGGCGAAGT 2242
Qy 3000 AATTCATCCAGATAGCAGTACGATTAATATATGGGC-GTCTCTAAAGATATATTACAA 3058
Db 2243 GCGGGGAGGATCTCTGTCTATCTCACCTTGTCTCTGCGGAGAAATATCCATCATGCG 2302
Qy 3059 TATCGGAGACACCGCAAGAACATTTGTTCTCCTCAATGACAGCCTGACACCGAAG 3118
Db 2303 TGATGCAATGCGCGGCTGCATACGCTTGATCCGGCTACCTGCCATTCGACCAACAGC 2362
Qy 3119 ACACGGGGTCTATTTTGTGCAAGCTTTACTTGGCTTCCCTGGTTGTCTTATGGG 3178
Db 2363 GAAC-----ATCGATCGAGCAGCAGTACTCGATGGAGCGGTCTTGTGCATCAGA 2419
Qy 3179 GCCAAGGACCCCGTCACTGCTCCAGTCTAAGCCACACGACGCGCAGCGCGGCTC 3238
Db 2420 TGATCTGGACGAAGACATCAGGGGTTCGCGCAGCGCAACTGTTTCGCGAGGCTCAAGC 2479
Qy 3239 CACCAACACCGCGGCCACCATCGGTTCGACGCCCTGCTCCCTGGCCCGCAGGGGCTC 3298
Db 2480 GCGCATGCCCAGCGGAGATCTCGTGTGATCCCATGGC----- 2519
Qy 3299 GGCACGCGCGGGGCGCAGTGCAACAGGGGGCTGGACTTCGCGCTGGATCCCAAC 3358
Db 2520 -----GATGCTGCT 2529
Qy 3359 TGTGTACTCTGTGATGAATCTCTTCACTATATGGTGTATTTCTACTGCTTGTTC 3418
Db 2530 TGCCGAATATCATGTGGAAATGCGCGCTTTCTGGATTCATCGACTGTG----- 2580
Qy 3419 TGAGATGAAGTTACGAGGCGGAGAGCCCCCGCTACGACGAGGCGCAGACCGC 3478
Db 2581 -----GCGGCTGGGTGTGGGAGCC 2601
Qy 3479 TCTATAACGAGCTCAATCTAGGACGAAGAGAGGATGATGTTTTGGACAAGAGACGTG 3538
Db 2602 GCTATCAGG----- 2610
Qy 3539 GCGGAGCCCTGAGATGGGGGGAAGCCGAGAAGGAACCCCTCAGGAAGGCGCTGTACA 3598
Db 2611 ----- 2610
Qy 3599 ATGACTGCAGAAAGATAGATGGGAGGCGCTACAGTGGATTTGGATGAAGGGGAGC 3658
Db 2611 -----ACATAGCGTGGCTACCCGTGATATTTCTGAAGAGCTTGGCG 2652
Qy 3659 GCGGAGGCGCAGGCGCAGTGGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACA 3718
Db 2653 GCGAATGGGCTGACCGCTCTCTGCTTTTACGGTATCGCGCTCCGATTCGACGGCA 2712
Qy 3719 CTTACGACCCCTTCACTGACGGCCCTGCCCCCTCGCTAACTCGACGCGCGCGGATC 3778
Db 2713 TCGCCTTCTATCGCCTTCTTGACG----- 2736

Qy 3779 CGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTGACTCA 3838
Db 2737 -----AGTTCTTCTGAGGGATCCGTCGACCTGCAGCCAAGCTT----- 2775
Qy 3839 ACAATATCACGAGCTGAAGCCTATAGATGACGAGCATAGATAAATAAAGATTTTAT 3898
Db 2776 -----ATCGATAAAATAAAGATTTTAT 2799
Qy 3899 TAGTCTCAGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAGCTAGCTTA 3958
Db 2800 TAGTCTCAGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTTA 2859
Qy 3959 AGTAACGCCATTTTCAAGGATGAAAATACATACTGAGAAATAGAGAAAGTTTCAGATCA 4018
Db 2860 AGTAACGCCATTTTCAAGGATGAAAATACATACTGAGAAATAGAGAAAGTTTCAGATCA 2919
Qy 4019 AGTTAGAAACAGAGAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTGAACAGTT 4078
Db 2920 AGTTAGAAACAGAGAGACAGC----- 2941
Qy 4079 CCTGCCCGCTCAGGGCCCAAGAACAGTTGGAAACAGAGAAATATGGGCCAAACAGGATATC 4138
Db 2942 -----AGAATATGGGCCAAACAGGATATC 2965
Qy 4139 TGTGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTGCCAGATCGGTC 4198
Db 2966 TGTGTAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTGCCAGATCGGTC 3025
Qy 4199 CGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTG-A 4257
Db 3026 CGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAA 3085
Qy 4258 AATGACCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGCTTCTGCTCGCG 4317
Db 3086 AATGACCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGCTTCTGCTCGCG 3145
Qy 4318 CTTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTC 4377
Db 3146 CTTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCTC 3205
Qy 4378 CGATAGACTGCTGCTCCCGGTPACCCGTGTTCTCAATAAACCCCTCTGAGTTGATCCG 4437
Db 3206 CGATAGACTGCTGCTCCCGGTPACCCGTGTTCTCAATAAACCCCTCTGAGTTGATCCG 3265
Qy 4438 ACTGCTGCTGCTGCTTCTTGGAGGGTCT-CTCTGAGTGAATGACTACCCGTCAGCG 4496
Db 3266 ACTTGTGCTGCTGCTTCTTGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCG 3325
Qy 4497 GGGTCTTTTCAAGTTTCTCCACCTACACAGGCTCTCACTAACTCTGATGTGCGCGAGG 4556
Db 3326 GGGTCTT----- 3333
Qy 4557 ACTCCGTCAGCCCGGTTTTTGTGTTATAATAAAGCAAGACAGTGTCCCTTCAAGCCA 4616
Db 3334 ----- 3333
Qy 4617 GACTACATCCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGC 4676
Db 3334 ----- 3333
Qy 4677 CACACGACTTTTAAAGTTTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCTCG 4736
Db 3334 ----- 3333
Qy 4737 TCTATTTTGTGGCTTCTCGTATTTTAAATTTTAGTTTGAATTCCTCCCTTCTGTGAGCA 4796
Db 3334 ----- 3333
Qy 4797 CGGCAATTCAGAGTAGTTAATCTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTCGGCG 4856
Db 3334 ----- 3333

QY 4857 TCAGTGTGAGATTGTCGCATAAAGGGGTCTGCCCTCTGTGTACAGACAGATCGGAATC 4916
Db 3334 -----TCAANN 3361
QY 4917 TAGAGTGCACTACAGATCCCGCGGTTCGGGGCTCTGATCTCAGGGCATCTTTGCT 4976
Db 3362 NNN 3421
QY 4977 AGAGATCCTCAGCGCGGACGATCGTGGCGGGTACGAGCTCGAATTCGTAATCATGG 5036
Db 3422 NNN 3481
QY 5037 TCATAGCTGTTCTCTGTGTGAAATGTTATCGCTCACAATTCACACAAATACGAGCC 5096
Db 3482 TCATAGCTGTTCTCTGTGTGAAATGTTATCGCTCACAATTCACACAAATACGAGCC 3541
QY 5097 GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAAGAGTAACTCAGATTAATTCG 5156
Db 3542 GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAAAGAGTAACTCAGATTAATTCG 3601
QY 5157 TTGGGCTCACTGCCCGCTTTCAGTCCGGAACCTGTCTGCCAGCTGCATTAATGAATC 5216
Db 3602 TTGGGCTCACTGCCCGCTTTCAGTCCGGAACCTGTCTGCCAGCTGCATTAATGAATC 3661
QY 5217 GGCCNAACGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGGCTTCCTCGCTCACT 5276
Db 3662 GGCCNAACGCGGGAGAGCGGTTTGGTATTTGGCGCTCTTCGGCTTCCTCGCTCACT 3721
QY 5277 GACTCGCTGCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGAGCGGTA 5336
Db 3722 GACTCGCTGCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGAGCGGTA 3781
QY 5337 ATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAATGTAGCAAAAGGCCAG 5396
Db 3782 ATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAATGTAGCAAAAGGCCAG 3841
QY 5397 CAAAAGGCGAGAAACCGTAAAGGCGGTTGCTGGGTTTTTCCATAGGCTCCGCC 5456
Db 3842 CAAAAGGCGAGAAACCGTAAAGGCGGTTGCTGGGTTTTTCCATAGGCTCCGCC 3901
QY 5457 CCTCAGCAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTA 5516
Db 3902 CCTCAGCAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTA 3961
QY 5517 TAAAGATACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCC 5576
Db 3962 TAAAGATACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTGTTCCGACCC 4021
QY 5577 CCGCTTACCGGATACCTGTCGCTTTCTCCCTCGGGAAGCGTGGCTTTCTCATAGC 5636
Db 4022 CCGCTTACCGGATACCTGTCGCTTTCTCCCTCGGGAAGCGTGGCTTTCTCATAGC 4081
QY 5637 TCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTCGTCTCAAGCTGGGCTGTGCAC 5696
Db 4082 TCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTCGTCTCAAGCTGGGCTGTGCAC 4141
QY 5697 GAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAAC 5756
Db 4142 GAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAAGTCAAC 4201
QY 5757 CCGGTAAAGACAGCTTATCCCACTGGCAGCAGCACTGTGTAAAGAGTATGAGAGCG 5816
Db 4202 CCGGTAAAGACAGCTTATCCCACTGGCAGCAGCACTGTGTAAAGAGTATGAGAGCG 4261
QY 5817 AGGTATGAGCGGTGTACAGATTTCTTGAAGTGGGCTTAACACTACGCTACACTAGA 5876
Db 4262 AGGTATGAGCGGTGTACAGATTTCTTGAAGTGGGCTTAACACTACGCTACACTAGA 4321
QY 5877 AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTCTCGGAAGAGGTTGGT 5936
Db 4322 AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTCTCGGAAGAGGTTGGT 4381
QY 5937 AGCTCTTGATCCGGCAAAACCAACCCGCTGTGTAGCGGTGTTTTTTTGTGCAAGCAG 5996

Db 4382 AGCTCTTGATCCGGCAAAACCAACCCGCTGTGTAGCGGTGTTTTTTGTTGCAAGCAG 4441
QY 5997 CAGATTACGCGCAGAAAAAGAGTCTCAAGAAATCTTTTGTATCTTTTCTACGGGCT 6056
Db 4442 CAGATTACGCGCAGAAAAAGAGTCTCAAGAAATCTTTTGTATCTTTTCTACGGGCT 4501
QY 6057 GAGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTTCATGAGATTAATCAAAAAG 6116
Db 4502 GAGCTCAGTGGAAACGAAAACTCAGTTAAGGGATTTTGGTTCATGAGATTAATCAAAAAG 4561
QY 6117 ATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAAT 6176
Db 4562 ATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAAT 4621
QY 6177 GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTTCTCAGCGATC 6236
Db 4622 GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGGCACTTCTCAGCGATC 4681
QY 6237 TGTCTATTTGCTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGATAAATCAGATACG 6296
Db 4682 TGTCTATTTGCTTCATCCATAGTTGCTGACTCCCGTCTGTGTAGATAAATCAGATACG 4741
QY 6297 GAGGGCTTACCATCTGGCCCCAGTCTGCATGATACCGCAGACCCACGCTCAGCGCT 6356
Db 4742 GAGGGCTTACCATCTGGCCCCAGTCTGCATGATACCGCAGACCCACGCTCAGCGCT 4801
QY 6357 CCAGATTATCAGCAATAAACACAGCCAGCGGAAGGCGGAGCGCAGAAAGTGGTCTGCA 6416
Db 4802 CCAGATTATCAGCAATAAACACAGCCAGCGGAAGGCGGAGCGCAGAAAGTGGTCTGCA 4861
QY 6417 ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCCGGGAAGCTAGAGTAAGTTCG 6476
Db 4862 ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCCGGGAAGCTAGAGTAAGTTCG 4921
QY 6477 CCAGTTAATAGTTTGGCGAAGCTTTGTCATGCTACAGCG - TCGTGGTGTACGCTCG 6535
Db 4922 CCAGTTAATAGTTTGGCGAAGCTTTGTCATGCTACAGCGCATCGTGGTGTACGCTCG 4981
QY 6536 TCGTTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACAGATCAAGCGAGTTACATGATCC 6595
Db 4982 TCGTTTGGTATGCTTCATTCAGCTCCGGTTCCTCAACAGATCAAGCGAGTTACATGATCC 5041
QY 6596 CCAATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCCGATCGTTGTGTCAGAAAGTAA 6655
Db 5042 CCAATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCCGATCGTTGTGTCAGAAAGTAA 5101
QY 6656 TTGGCCGCGAGTTTATCATCTAGGTTATGGCAGCAGCTGATTAATCTCTTACTGTCTATG 6715
Db 5102 TTGGCCGCGAGTTTATCATCTAGGTTATGGCAGCAGCTGATTAATCTCTTACTGTCTATG 5161
QY 6716 CCATCCGTAAGATGCTTTTCTGCTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAG 6775
Db 5162 CCATCCGTAAGATGCTTTTCTGCTGACTGGTGAGTACTCAACCAAGTCAATCTGAGAAATAG 5221
QY 6776 TGTATGCGGGACCGAGTTGCTCTTTGCCCGCGCTCAATACGGGATAATACCGCGCACAT 6835
Db 5222 TGTATGCGGGACCGAGTTGCTCTTTGCCCGCGCTCAATACGGGATAATACCGCGCACAT 5281
QY 6836 AGCAGAACTTTTAAAGTGTCTCATCAATGGAAAAAGCTTTCTCGGGGCGAAAACTCTCAAG 6895
Db 5282 AGCAGAACTTTTAAAGTGTCTCATCAATGGAAAAAGCTTTCTCGGGGCGAAAACTCTCAAG 5341
QY 6896 ATCTTACCGCTGTGAGATCCAGTTCCGATTAACCCACTCGTGCACCACTGATCTTCA 6955
Db 5342 ATCTTACCGCTGTGAGATCCAGTTCCGATTAACCCACTCGTGCACCACTGATCTTCA 5401
QY 6956 GCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCA 7015
Db 5402 GCATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCGCA 5461
QY 7016 AAAAAAGGAATAAGGGCGACACGGAATGTTGAATCTCATACTCTCTCTTTTCAATPAT 7075

1686 ATCGCTCACACCAAGTCGCTAGATGTCAAGAGAGAGCGTTGGGTTACTCTTCTGCTCTGCA 1745
1305 ATCCGCTCACACCAAGTCGCTAGATGTCAAGAGAGAGCGTTGGGTTACTCTTCTGCTCTGCA 1364
1746 GAATGGCCAACTTTAAACGTCGATGGCCGCGAGACGCGACCTTTAAACCGAGACCTCATC 1805
1365 GAATGGCCAACTTTAAACGTCGATGGCCGCGAGACGCGACCTTTAAACCGAGACCTCATC 1424
1806 ACCAGGTTAAGATCAAGGTTCTTTTCACTGGCCCGGATGAGACACCCAGACCGAGTCCCC 1865
1425 ACCAGGTTAAGATCAAGGTTCTTTTCACTGGCCCGGATGAGACACCCAGACCGAGTCCCC 1484
1866 TACATCTGACCTGGGAGCCTTGGCTTTTGAACCCCTCCCTGGGGTCAAGCCCTTTGTA 1925
1485 TACATCTGACCTGGGAGCCTTGGCTTTTGAACCCCTCCCTGGGGTCAAGCCCTTTGTA 1544
1926 CACCTAAGCCTCCGCTCTCTTCTCCATCCGCGCGCTCTCCCTCTTGAACCTCCT 1985
1545 CACCTAAGCCTCCGCTCTCTTCTCCATCCGCGCGCTCTCCCTCTTGAACCTCCT 1604
1986 GGTTCGACCCGCTCGATCTCTCTTATCCAGCCCTCACTCTCTCTAGGCGCCCC 2045
1605 CGTTGACCCGCTCGATCTCTCTTATCCAGCCCTCACTCTCTTAGGCGCCCC 1664
2046 ATATGGCCATATGATCTTATATGGGCGACCCCGCCCTTGTAACTTCCCTGACCT 2105
1665 ATATGGCCATATGATCTTATATGGGCGACCCCGCCCTTGTAACTTCCCTGACCT 1724
2106 GACATGACAGAGTTACTAACAGCCCTCTCTCAAGCTCACTTACAGGCT-TCTACTTA 2164
1725 GACATGACAGAGTTACTAACAGCCCTCTCTCAAGCTCACTTACAGGCTCTCTACTTA 1784
2165 GTCCAGCAGAAAGTCTGGAGACCTCTGGCGGACCTTACCAAGAACAACTGGACCGACCG 2224
1785 GTCCAGCAGAAAGTCTGGAGACCTCTGGCGGCA-CGTACCAAGAACAACTGGACCGACCG 1843
2225 GTGGTACTCACCCTTACCGAGTGGCGGACACAGTGTGGTGGCGGACACAGCTAAG 2284
1844 GTGGTACTCACCCTTACCGAGTGGCGGACACAGTGTGGTGGCGGACACAGCTAAG 1903
2285 AACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTGCTGACACCCCGCCCGCCCTC 2344
1904 AACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTGCTGACACCCCGCCCGCCCTC 1963
2345 AAAGTAGACGGCATCGCAGCTTGGATACAGCCGCCACAGTGGAGGCTGCGACCCCGGG 2404
1964 AAAGTAGACGGCATCGCAGCTTGGATACAGCCGCCACAGTGGAGGCTGCGACCCCGGG 2023
2405 GGTGGACCATCTCTAGACTGCGGATGGGATGGAGCTGTATCATCTCTTCTTGTGTAGCAA 2464
2024 GGTGGACCATCTCTAGACTGCGGATGGGATGGGATGGAGCTGTATCATCTCTTCTTGTGTAG 2082
2465 CAGCTACAGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGGCCA 2524
2083 CGGCCGCGAGTGTCTGGAGCCCATCTCTATTATAAAAA-----TGCTCTAGAGCAA 2134
2525 GCSTGGGTGACAGAGTGACCATCACCTGTAGGCGCAGTCAGGATGTGGGTACTTCTGTAG 2584
2135 CCGGGAGCTGGTGTGTGATCTCTCTCTCTACAGCTCTCCAGAAAGGATACAGCTGGAG 2194
2585 CTTGGTTACAGAGCAAGCCAGGTTAAGGCTCCAAAGCTGTGTATCTACTGACATPCACCC 2644
2195 TCAGTTTTAGCGATGTCGAAGAGAACAGAGCTGGAAGCCCGAGAGAAACTGAACAGAAAG 2254
2645 GGCACATGTTGTCGCAAGAGATTCAGCGGTAGCGGTAGCGGTACCGACTTACCTTCA 2704
2255 GGAGACCCCGAGTGCCATCAATGGAACCCATCTCTGGCAC--TGGCGGATAGCCCGCG 2312
2705 CCATCAGCAGCCTCCAGCAGAGGACATCGCCACTACTGTCCAGCAATATAGCCTCT 2764
2313 GTGNATGGAGCCACTGGCCACAG-----CAGCAGTTTGGATGCG 2351
2765 ATCGGTCGTTGGCCAAAGGACCAAGGTGGAAATCAAAACGAGGTGGCTCAGGATCGGGTG 2824

2352 CGGAGGTAATCCCTCCATGGCAGCAGTGAAGCAAGCGCTGAGAGAGGCTTGGCGAT----- 2405
2825 GATCCGGCTCTGTTGGCTCAGGATCGAGGTCCTAACTGGTGGAGAGCGGTGGAGGTGTTG 2884
2406 -----GAGTTTGAATCGGTACCGGAGAGCAATTCAGTG 2439
2885 TGAACCTGGCCGGTCTCCCTCGGCTGTCTGCTCGCATCTGGCTTCGATTTCACCAT 2944
2440 ATCTAACATCCCAG-----CTTCAT 2459
2945 ATTGGATGAGTGGGTGAGACGACCTCGAAAGGCTTGTAGTGGATTTCGAGAAATTC 3004
2460 ATAACCCAGGACAGTATATCAGAGCTTTTGAACAGGTAGTG----- 2501
3005 ATCCAGATAGCAGTACGATTAATCTATGCGCGCTCTCTAAAGGATAGATTTACATATCGC 3064
2502 ----- 2501
3065 GAGACAAAGCCAAAGAACACATTTGTTCTGCAATGGAACGCTGAGACCCGAGACACCG 3124
2502 -----AATGAACTCTTTGGGATGGGTAAACTG 2530
3125 GGGTCTATTGTTGCAAGCCTTTACTTTGCGCTTCCCTGCTTCTTATTGGGGCCAAG 3184
2531 GGGTGGCATTTG----- 2542
3185 GGACCCCGGTCAACGCTCTCCAGTGTAAAGCCACACGACGCGCAGCCGCGACCAACAA 3244
2543 ----- 2542
3245 CACCGGCGCCACATCGCGTCGAGCCCTGTCTCTGCGCCAGAGCGGCTCGGCCAG 3304
2543 -----GGCTCTCTCTCTTTGGGGGGCACTGTGCGTGGAAAGCGTAGACAAAGAGATGC 2598
3305 CGCGGGGGCGCAGTGCACACGAGGGGCTGCACTTCGCGCTGGATCCCAAACTCTGCT 3364
2599 AGGATTTGGTGAATGCGGATTTGCAAGTTGGATGGCCACTTACCTGAATGACCACTAGAGC 2658
3365 ACCTGTGGAATGGAATCTCTTCATCTATGCTGTCTCTCACTGCTCTTCTCTGAGAG 3424
2659 CTTGGATCCAGGAGAACGGC-----GGCTGGGACACTTTTG 2694
3425 TGAAGTTTACGAGAGCGCAGAGCCCGCCGTAACAGAGGCGCAGAAACAGGCTTATA 3484
2695 TGGATCTCTACGGGAAACAATACAGACACCCGAGAGCGGAAAGCGCAGAGCGTTTCAACC 2754
3485 ACAGCTCAATCTTAGGACGAGAGAGAGTACGATGTTTTTGGACAAGAGAGTGGCCGG 3544
2755 GCTGGTTCTCTGACGGGCATGACTGTGGCTGGTGTAGTTCTG----- 2795
3545 ACCTGAGATGGGGGAAAGCCGAGAAAGAACCCCTCAGGAAGGCGCTGTACATGAAC 3604
2796 ----- 2795
3605 TGCAGAAAGATAAGATGGCGAGGCGCTTACAGTGAATTTGGGATGAAGGCGAGCCGGA 3664
2796 ----- 2795
3665 GGGGCAAGGGGACGATGGCCTTTACAGGGTCTCAGTACAGCCACCAAGACACCTACG 3724
2796 ----- 2795
3725 ACGCCCTTACATGACAGGCCCTGCCCTCGCTAACTCGACGCGCGCGGATCCGGATT 3784
2796 -----CTGGGCTCACTCTTCACTCGGAAGTGAACAGATT 2829
3785 AGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGACTCAACAATA 3844
2830 CTGAGATATCCATCACACTGGCGGGAATTTCCGGGGCGCGATCCGAGTTCTTCTGA 2889
3845 TCACAGCTGAAGCCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTAGTCT 3904


```
Qy 6061 CTCAGTGAACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 6120
Db 4702 CTCAGTGAACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 4761
Qy 6121 TCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGT 6180
Db 4762 TCACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATAGT 4821
Qy 6181 AAATCTGGTGTGACAGTTTACAAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTC 6240
Db 4822 AAATCTGGTGTGACAGTTTACAAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCTGTC 4881
Qy 6241 TATTTCTGTTTCATCATAGTTTCCCTGACTCCCGGCTGCTAGATACTACGATCGGGAGG 6300
Db 4882 TATTTCTGTTTCATCATAGTTTCCCTGACTCCCGGCTGCTAGATACTACGATCGGGAGG 4941
Qy 6301 GCTTACCATCTGGCCCAAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360
Db 4942 GCTTACCATCTGGCCCAAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 5001
Qy 6361 ATTTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGCAGAGTGGTCTCTGCAACTT 6420
Db 5002 ATTTATCAGCAATAAACCAGCAGCCGGAAGGGCCGAGCGCAGAGTGGTCTCTGCAACTT 5061
Qy 6421 TATCGGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTTCGGCAG 6480
Db 5062 TATCGGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTTCGGCAG 5121
Qy 6481 TTAATAGTTTCGCAAGCTGCTGCCATTTGCTCAGGCAATGCTGCTGCTGCTGCTGCT 6539
Db 5122 TTAATAGTTTCGCAAGCTGCTGCCATTTGCTCAGGCAATGCTGCTGCTGCTGCTGCTGCT 5181
Qy 6540 TTGTTATGGCTTCATTCAGCTCCGGTTCCTCAACAGTCAAGCGGAGTACATGATCCGCCA 6599
Db 5182 TTGTTATGGCTTCATTCAGCTCCGGTTCCTCAACAGTCAAGCGGAGTACATGATCCGCCA 5241
Qy 6600 TGTGTGCAAAAAGCGGTTAGCTCTTCGCTGCTCCGATCGTTGTCGAGAAGTAAAGTGG 6659
Db 5242 TGTGTGCAAAAAGCGGTTAGCTCTTCGCTGCTCCGATCGTTGTCGAGAAGTAAAGTGG 5301
Qy 6660 CCGCAGTGTATCACTCATGTTTATGCGAGCACTGCATTAATCTCTTACTGTCTATGCCAT 6719
Db 5302 CCGCAGTGTATCACTCATGTTTATGCGAGCACTGCATTAATCTCTTACTGTCTATGCCAT 5361
Qy 6720 CCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTTGAGAAGTGTGA 6779
Db 5362 CCGTAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTTGAGAAGTGTGA 5421
Qy 6780 TGGCGGACCGAGTTGCTTTGCGCGGCTCAATACGCGGATAATACCGCGCCACATAGCA 6839
Db 5422 TGGCGGACCGAGTTGCTTTGCGCGGCTCAACACGCGGATAATACCGCGCCACATAGCA 5481
Qy 6840 GAACTTTTAAAGTGTCTATGGAATAAGCTTTCTTGGGGGGAATACTCTCAAGGATCT 6899
Db 5482 GAACTTTTAAAGTGTCTATGGAATAAGCTTTCTTGGGGGGAATACTCTCAAGGATCT 5541
Qy 6900 TACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCTGTCACCACTGATCTTTTCAAGAT 6959
Db 5542 TACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCTGTCACCACTGATCTTTTCAAGAT 5601
Qy 6960 CTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAA 7019
Db 5602 CTTTTACTTTTCCACGCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAA 5661
Qy 7020 AGGGAATAAGGGGACACGGAATTTGGAATCACTACTCTTCTGTTTCAATTAATTT 7079
Db 5662 AGGGAATAAGGGGACACGGAATTTGGAATCACTACTCTTCTGTTTCAATTAATTT 5721
Qy 7080 GAAGCATTTATCAGGTTTATGTTCTCATGACGGATACATATTTGAATGATTTTAGAAAA 7139
Db 5722 GAAGCATTTATCAGGTTTATGTTCTCATGACGGATACATATTTGAATGATTTTAGAAAA 5781
Qy 7140 ATAAACAAATAGGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGCTTAAGAAA 7199
```

```
Db 5782 ATAAACAAATAGGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAA 5841
Qy 7200 CCATTATTATCATGACATTAACCTATAAAATAGCGGTATCACGAGGCCCTTTTGGTCT 7257
Db 5842 CCATTATTATCATGACATTAACCTATAAAATAGCGGTATCACGAGGCCCTTTTGGTCT 5899

RESULT 6
AX491314 LOCUS AX491314 5782 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 1 from Patent WO0234929.
ACCESSION AX491314
VERSION AX491314.1 GI:22324009
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kohn, B.Y. and Daley, G.O.
TITLE Expression vectors and uses thereof
JOURNAL Patent: WO 0234929-A 1 02-MAY-2002;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES
source Location/Qualifiers
1..5782
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Synthetically generated nucleic acid"

ORIGIN
Query Match 42.0%; Score 3213; DB 6; Length 5782;
Best Local Similarity 73.3%; Pred. No. 0;
Matches 4860; Conservative 0; Mismatches 905; Indels 864; Gaps 24;
Qy 552 TGAAGAGCCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGCAGGCAAT 611
Db 3 TGAAGAGCCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGCAGGCAAT 62
Qy 612 GG-AAAAATACATACTCAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAACAGA-GAGACAG 669
Db 63 GGAATAATACATACTCAGAAATAGAGAAAGTTTCAGATCAAGGTCAGGAACAGATGGAACAG 122
Qy 670 CAGAAATATGGGCCCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCC- GTCAGGGCCAA 728
Db 123 CTGAATATGGGCCCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCCCGCTCAGGGCCAA 182
Qy 729 GAACAGTTGGAACAGGAAGATATGGGCCCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCC 788
Db 183 GAACAGATGGAACAGGCTGAATATGGGCCCAACAGGATATCTGTGTAAAGCAGTTCTTCCGCC 242
Qy 789 CCGGCTCAGGGCCAAAGAACAGATGCTGCCAGATGCGGTCCCGCTCCCGCTCAGCAGTTTCTAG 848
Db 243 CCGGCTCAGGGCCAAAGAACAGATGCTGCCAGATGCGGTCCCGCTCAGCAGTTTCTAG 302
Qy 849 AGAACCATCAGATGTTTTCCAGGGTCCCGCAAGGACCTGAAATAGACCTTGTCCTTATTTG 908
Db 303 AGAACCATCAGATGTTTTCCAGGGTCCCGCAAGGACCTGAAATAGACCTTGTCCTTATTTG 362
Qy 909 AACTAAACCAATCAGTTCTGCTTCTGCTTCTGTTTCCGGGCTTCTGCTCCCGCAGTCAAT 968
Db 363 AACTAAACCAATCAGTTCTGCTTCTGCTTCTGTTTCCGGGCTTCTGCTCCCGCAGTCAAT 422
Qy 969 AAAGAGCCACACACCTCTCATCTCGGCGCGCAGTCTCCGATAGACTGCGTCCGCGGG 1028
Db 423 AAAGAGCCACACACCTCTCATCTCGGCGCGCAGTCTCCGATAGACTGCGTCCGCGGG 482
Qy 1029 TACCCGTTATTTCCCAATAAAGCCTCTTGTGTTTTCATCCGAAATCGTGACTCGTGATCC 1088
Db 483 TACCCGTTATTTCCCAATAAAGCCTCTTGTGTTTTCATCCGAAATCGTGACTCGTGATCC 541
Qy 1089 TTGGGAGGGTCTCTCTCAGATTGATGTCGCC- ACCTCGGGGGTCTTTTCAATTTGGAGGT 1147
```

Db 542 TTGGAGGGCTCTCTCTGAGTGATTGACTACCGTTCAGCGGGGCTCTTTCAATTGGGGGC 601
 QY 1148 TCACCGAGATTGGAGACCCCTGCCACGAGCCACCGACCCCGCGGGAGGTAAAGC 1207
 Db 602 TGTCCGGGATCGGGAGACCCCTGCCAGAGGACCCGA-CCACACCCCGGGAGGTAAAGC 660
 QY 1208 TGCCAGCAACTTATCTGTGTCTGTCCGATTTGTAGTGTCTATGACTGATTTATGCGC 1267
 Db 661 TGCCAGCAACTTATCTGTGTCTGTCCGATTTGTAGTGTCTATGACTGATTTATGCGC 720
 QY 1268 CTGCGTCTGATCTAGTTAGTAACTAGTCTGTATCTGCGGACCGGTGTAAGTAACTGAC 1327
 Db 721 CTGCGTCTGATCTAGTTAGTAACTAGTCTGTATCTGCGGACCGGTGTAAGTAACTGAC 780
 QY 1328 GAGTTCCGGAACACCCCGCGCACCTTGGAGACGTCCTCAGGAGCTTTCGGGGGCGGTTTT 1387
 Db 781 GAGTTCCGGAACACCCCGCGCACCTTGGAGACGTCCTCAGG-TCGGGGGCGGTTTT 836
 QY 1388 TGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTGGACTCTTTGGTGCAACCCCGCTT 1447
 Db 837 TGTGGCCCGACCTGAGTCTCAAAATCCCGATCGTTTGGACTCTTTGGTGCAACCCCGCTT 896
 QY 1448 AGAGGAGGATATGTGGTTCTGTGTAGGAGACGAGAACTTAAACAGTTCCCGCCTCCGTC 1507
 Db 897 AGAGGAGGATATGTGGTTCTGTGTAGGAGACGAGAACTTAAACAGTTCCCGCCTCCGTC 956
 QY 1508 TGAATTTTTCGTTTTCGTTTGGAGCCGAGCCGCGCGCGTCTTGTCTGTCGACGAT 1567
 Db 957 TGAATTTTTCGTTTTCGTTTGGAGCCGAGCCGCGCGCGTCTTGTCTGTCGACGAT 1016
 QY 1568 CGTTCTGTGTCTCTGTCTGACTGTGTCTGTATTTGTCTGAAATATGGGCCCGG 1627
 Db 1017 CGTTCTGTGTCTCTGTCTGACTGTGTCTGTATTTGTCTGAAATATGGGCCCGG 1076
 QY 1628 CTAGACTGTACACCTCCCTTAAAGTTTGAACCTTGTAGTCTACCTGAAAGATGTCCAGCGAT 1687
 Db 1077 CCAGACTGTACACCTCCCTTAAAGTTTGAACCTTGTAGTCTACCTGAAAGATGTCCAGCGAT 1136
 QY 1688 CGCTCAACCAAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACCTTCTGCTCTGCAGA 1747
 Db 1137 CGCTCAACCAAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACCTTCTGCTCTGCAGA 1196
 QY 1748 ATGGCCAACTTTAACTCGGATGGCGCGAGACGCACTTTAAACCGAGACTCATCAC 1807
 Db 1197 ATGGCCAACTTTAACTCGGATGGCGCGAGACGCACTTTAAACCGAGACTCATCAC 1256
 QY 1808 CCAGGTTAAGATCAAGGTCTTTTACCTGCGCGCATGACACCCAGACAGGTCCCTTA 1867
 Db 1257 CCAGGTTAAGATCAAGGTCTTTTACCTGCGCGCATGACACCCAGACAGGTCCCTTA 1316
 QY 1868 CATCTGACCTGGGAAGCTTGGCTTTTGAACCCCTCTCTGGGTCAAGCCCTTTGTACA 1927
 Db 1317 CATCTGACCTGGGAAGCTTGGCTTTTGAACCCCTCTCTGGGTCAAGCCCTTTGTACA 1376
 QY 1928 CCTAAGCTCGGCTCTCTCTCTCATCTCGGCGGCTCTCTCCCGCTTGAACCTCTCG 1987
 Db 1377 CCTAAGCTCGGCTCTCTCTCTCATCTCGGCGGCTCTCTCCCGCTTGAACCTCTCG 1436
 QY 1988 TTGACCCCGCTCTGATCTCTCTCTTATCCAGCCCTCACTCTCTCTAGGCGCCCAT 2047
 Db 1437 TTGACCCCGCTCTGATCTCTCTCTTATCCAGCCCTCACTCTCTCTAGGCGCCCAT 1496
 QY 2048 ATGGCCATATGATCTTATATGGGACCCCGCGCCCTTGTAACTTCTCTGACCCCTGA 2107
 Db 1497 ATGGCCATATGATCTTATATGGGACCCCGCGCCCTTGTAACTTCTCTGACCCCTGA 1556
 QY 2108 CATGAAGAGTTTACTTAAAGCCCTCTCTTCAAGCTCACTTACAGGCT-TCCTACTAGT 2166
 Db 1557 CATGAAGAGTTTACTTAAAGCCCTCTCTTCAAGCTCACTTACAGGCTCTCTACTTAGT 1616
 QY 2167 CCAGCAAGAGTCTGAGACCTCTGCGGCGAGCTTACCAAGAACAACTGAGACCGAGT 2226
 Db 1617 CCAGCAAGAGTCTGAGACCTCTGCGGCGAGCTTACCAAGAACAACTGAGACCGAGT 1676

QY 2227 GGTACTCACTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACAGACTAAGAA 2286
 Db 1677 GGTACTCACTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACAGACTAAGAA 1736
 QY 2287 CCTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGTGACCAACCCCGCCTCAA 2346
 Db 1737 CCTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGTGACCAACCCCGCCTCAA 1796
 QY 2347 AGTAGACGCGATCGAGCTTGGATACACGCGCCACAGTGAAGGCTGCCGACCCCGGGG 2406
 Db 1797 AGTAGACGCGATCGAGCTTGGATACACGCGCCACAGTGAAGGCTGCCGACCCCGGGG 1856
 QY 2407 TGACCATCTCTAGACTTGCATGGATGGAGCTGTATCATCTCTTCTGTGTAGCAACA 2466
 Db 1857 TGACCATCTCTAGACTTGCATGGATGGAGCTGTGTGTGTAGGGAATTAAGCTTGTATCTC 1916
 QY 2467 GCTACAGGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGCGCACG 2526
 Db 1917 TATAATCTCGGCAACCTATTTTCCCTCGAACACCTTTTAAAGCGTAGATAAAGGCT 1976
 QY 2527 GTGGGTGACAGAGTGAACCATCACTGTAAAGCCAGTCAAGATGTGGGTACTTCTGTAGCT 2586
 Db 1977 GGGACACTTTCATGAGCGAAATAATACATCGTCACTGGGACATGTGACAGATCCATGC 2036
 QY 2587 TGGTACCAGCAGACGAGTAAAGCTCCAAAGCTGTCTACTTGGACATCCACCCGG 2646
 Db 2037 AGCTAACTCGCAACCGACTGATG-CTTCTGAAACAATGGAAGGCAATATTG 2089
 QY 2647 CACACTGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACTTCACTC 2706
 Db 2090 CCGTAAGCGTGGCGGTCTGTGTACCGTGGGTGAAGACCCAGAAACAGCACTCGATC- 2146
 QY 2707 ATCAGAGCTCCAGCAGAGACATCGCCACTACTACTGACAGCAATATAGCTCTAT 2766
 Db 2147 -TGAGCGCGATTTTCCCGAGCGTTTCAACGCGCTGTATGGCGAGATCGATCCCGTC- 2202
 QY 2767 CGGTCTTCCGCAAGGACCAAGGTGGAATCAAAACGAGGTGGCTCAGGATCGGTGA 2826
 Db 2203 -----GTTTTACAAGTCTGTGACTGGGAAACCCCTGGGTTA 2239
 QY 2827 TCCTGCTGTGGTTCAGGATCGGAGGTCCAACCTGGTGGAGCGGTGGAGGTGTGTG 2886
 Db 2240 CCAACTTAATGCGCTTGGAGACATCCCGCTTTCGCCAGTGGCGTAAATAGCAAGAG 2299
 QY 2887 CAACTGGCGGTCCCTGGCGCTGTCTGCTCGCATCTGGCTTCGATTTTACCACATAT 2946
 Db 2300 CCGGCACCGATCGCGCTTCCCAACAGTTTCCGACGCTGAAATGGCGAATGGCGCTTGC 2359
 QY 2947 TGGATGAGTTGGGTGAGACAGCACCTTGAAAGGTCTTGGATGGATTCGAGAAATTCAT 3006
 Db 2360 TGGTTTCGGCACCAAGACGGTGGCGGAAAGCTGGTGGATGGAT----- 2407
 QY 3007 CCAGATAGCAGTACGATTAACCTATGCGCGCTCTCTAAAGAGTAGATTTACAATATCGCA 3066
 Db 2408 --CTTCTGAGCGCGATCTGTCTGCTCCCTCAAACTGGCAGATGACGTTACGATG 2465
 QY 3067 GACAAAGCAGAAACATTTGTTCTGCAAAATGGACAGCTGAGACCCGAGACACCGGG 3126
 Db 2466 CGCCCATCTACACCAACGAGCTATCCCATTAACGCTCAATCCGCGCTTGTTCACCG 2525
 QY 3127 GTCTATTTTGTGCAAGCCTTACTTTCGCTTCCCTGGTGTCTTATTTGGGCGCAAGGG 3186
 Db 2526 AGAATCCGAGCGGTGTACTCGCTCAATTTTATGTGTGTAAGTGGCTGGCTACAGGAA 2585
 QY 3187 ACCCGGCTCACCGTCTCCAGTGTCTAAGCCCAACAGCAGCCAGCGCGCACCAACA 3246
 Db 2586 GGCAGACCGCAATTTATTTTGTGCGGTAACTCGGCGTTTCACTG----- 2633
 QY 3247 CCGCGGCCCAACCATCGCTGCGAGCCCTGTCTCTGCGCCAGAGCGGCTCGGCCAGCG 3306
 Db 2634 -----TGGTGCAACGGGCGCTGGGTCTGTTACGGGCAAGACAGTCTG 2674

QY 3307 GCGGGGCGCAGTGCACACGAGGGGCTGGACTTCGCCCTGGATCCCAAACTCTGTAC 3366
Db 2675 TTTGGCGCTTAATTTGAGCTCGAGCGCATATCTACGCGCGGA-----GAAACCGCCTC 2730
QY 3367 CTGCTGGATGAACTCTTCACTATAGTGTCACTTCTCACTGCCTTGTCTCTGAGAGTG 3426
Db 2731 GCGGTGATGTGCTCGCTGAGTGACGCGGAGTTACT----- 2768
QY 3427 AAGTTACAGAGGCGCAGAGCCCGCTGACACAGCAGGGCCAGAACCACTCTATTAAC 3486
Db 2769 -----TGAAGATCAAGATATGTGGCGGATGAGCGGATTCGAGCGCAAAAC 2814
QY 3487 GAGCTCAATCTAGGACGAAGAGGAGTACGATGTTTTGGACAGAGAGTGGCGGGAC 3546
Db 2815 GGTCTGCGCTCGGGACGCGGAATTGAATTATGGCCCAACACAGAGTGGCGCGCGGAC 2874
QY 3547 CCTGAGATGGGGGAAAGCCAGAGGAAAGAACCTCAGGAAGGCGCTGTACAATGAACCTG 3606
Db 2875 TTCCAGTT-----CAACATCAGCGCTACAGTCAACAGCACTGATGGAAACCAAGCAT 2928
QY 3607 CAGAAAGATAAGATGGCGGAGCCCTACAGTGAGATTTGGGATGAAAGCGAGCGCGGAGG 3666
Db 2929 CGCCATCTGCTGCACGCGGAAGAACCGACATGGCTGTTATACGACGGTTTCCATATGGG 2988
QY 3667 GGCAAGGGCAGATGGCTTTACAGGGTCTCAGTACAGCCACCAAGGACACCTACGAC 3726
Db 2989 ATTGGTGGCGACGACTCCTGGAGCCCGTCACTGTGCGGGAATTCAGCTGAGCGCGGT 3048
QY 3727 GCCCTTACATGACAGGCCCTCGCCCTCGCTAACTCGACGCGCGCGGATCCGGATTAG 3786
Db 3049 CGCTACCATTTACAGTTGGTCTGGTGTCAAAATAAATAAATACCGGCGAGCCATGTCTG 3108
QY 3787 TCCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTGACTCAACAATATC 3846
Db 3109 CCGCTATTTCGTGAAGGAATCCATTATGT-----ACTATTGA 3147
QY 3847 ACCAGCTGAACCTTATAGAGTACGAGCCATAGATAAATAAAGATTTATTTAGTCTCC 3906
Db 3148 AACTCGAGCGCGCGCAGCAGCAGTGGTCGACGATAAATAAAGATTTATTTAGTCTCC 3207
QY 3907 AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTAAGTAAGC 3966
Db 3208 AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTAAGTAAGC 3267
QY 3967 CATTTTGAAGGCATGG-----AAAATACATACTGAGATAGAGAAGTTCAAGTCAAGGTTAG 4025
Db 3268 CATTTTGAAGGCATGGAAAATAACATACTGAGATAGAGAAGTTCAAGTCAAGGTCAG 3327
QY 4026 GAAACAGA-GAGACAGCAGATATGGGCCAAACAGAGATCTGTGTAGCAGTTCCTGCC 4084
Db 3328 GAAACAGATGGAACAGCTGAATATGGGCCAAACAGAGATCTGTGTAGCAGTTCCTGCC 3387
QY 4085 CC-GCTCAGGCGCAGACAGTTCGACAGAGATATGGGCCAAACAGAGATCTGTGG 4143
Db 3388 CCGCTCAGGCGCAGAAACAGATGGAACAGCTGAATATGGGCCAAACAGAGATATCTGTGG 3447
QY 4144 TAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCTGCC 4203
Db 3448 TAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCTGCC 3507
QY 4204 CTCAGCAGTTCTTATAGAACCAATCAGATGTTTTCAGGGTGGCCCAAGACCTGAAATGAC 4263
Db 3508 CTCAGCAGTTTCTAGAGAACCAATCAGATGTTTTCAGGGTGGCCCAAGACCTGAAATGAC 3567
QY 4264 CCTGCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTCTG 4323
Db 3568 CCTGCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTCTG 3627
QY 4324 CTCGCCGAGCTCAATAAAGAGCCCAACAAACCCCTCAGTTCGCGCGCGCTCCTCCGATAG 4383
Db 3628 CTCGCCGAGCTCAATAAAGAGCCCAACAAACCCCTCAGTTCGCGCGCGCTCCTCCGATG 3687
QY 4384 ACTCGTGGCGCGGTACCCGTTCTCAATAAACCCTCTTTGCAAGTTGCATCCGACTCGT 4443

Db 3688 ACTGAGTCGCCCGGTACCCGTGTATCAATAAAACCCCTTGTGCAGTTGCATCCGACTTGT 3747
QY 4444 GGTCTCGCTGTTCTTGGAGGGTCT-CTCTGAGTGATTGACTACCCGTCAGCGGGTCT 4502
Db 3748 GGTCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCCGTCAGCGGG--- 3804
QY 4503 TTCAGTTTCTCCCACTACACAGGCTCAACAGGCTCAATACTCTGATGTGCGCAGGGAATCGG 4562
Db 3805 ----- 3804
QY 4563 TCAGCCCGTTTTTGTGTTTATAAATAAATGCAAGACAGTGTTCCTTCAAGCCAGACTAC 4622
Db 3805 ----- 3804
QY 4623 ATCTGACTCTCGGCTTTATAAAAGATGTTGAAGGGCTCTGTGAGCTATCTGCCACACG 4682
Db 3805 ----- 3804
QY 4683 ACTTTTAAAGATTTTATGCTCTGATGAGGATTTAGTCAATCTATCTCTGTATTT 4742
Db 3805 ----- 3804
QY 4743 TTGCTGGCTTCTCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGACACGCGA 4802
Db 3805 ----- 3804
QY 4803 TTGCAGAGTAGTTAAATACTCTGAGGGCAGGCTTCTGTGAAAAGGTGCTCGGCTCAGTG 4862
Db 3805 ----- 3804
QY 4863 TGAGATTTTGCCATAAAGGGGTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGAGT 4922
Db 3805 ----- 3804
QY 4923 GCATACTCAGAGTCCCGCGGTCCGGGGTCTGTATCTCAGGGCATCTTTGCCGTAGAGAT 4982
Db 3805 ----- 3804
QY 4983 CCTCTACGCGGACGCATCGTGGCGGGTACCGAGTCCGAATTCGTAATCATGTGTCATAG 5042
Db 3805 ----- 3804
QY 5043 CTGTTTCTGTGTGAAATTTGTTATPCGCTCAAAATCCACACAACATACAGCAGCGGAAGC 5102
Db 3805 ----- 3804
QY 5103 ATAAAGTGAAGCCTGGGGTGCCCTAATGATGAGTAACTCACATTAATTTGCGTTGCGC 5162
Db 3805 ----- 3804
QY 5163 TCACTGCCCGTTTCCAGTCCGGAAACCTGTGTGCCAGCTGCATTAATGAATCGGCCAA 5222
Db 3805 -----GTCTTTTCATTCGCAATTAATGAATCGGCCAA 3835
QY 5223 CGCGCGGAGAGCGGTTTGGGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCG 5282
Db 3836 CGCGCGGAGAGCGGTTTGGGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACTCG 3895
QY 5283 CTGCGCTCGGTCTGTTGGCTGCGCGAGCGGTATCAGTCACTCAAAAGCGGTATACGG 5342
Db 3896 CTGCGCTCGGTCTGTTGGCTGCGCGAGCGGTATCAGTCACTCAAAAGCGGTATACGG 3955
QY 5343 TTATCCACAGAACTCAGGGGATACCGAGGAAGAACATGTGAGCAAAAGCCAGCAAAAG 5402
Db 3956 TTATCCACAGAACTCAGGGGATACCGAGGAAGAACATGTGAGCAAAAGCCAGCAAAAG 4015
QY 5403 GCCAGGAACCGTAAAAAGGCGCGTGTCTGCGGTTTTTCCATAGGCTCCGCCCCCTGAC 5462
Db 4016 GCCAGGAACCGTAAAAAGGCGCGTGTCTGCGGTTTTTCCATAGGCTCCGCCCCCTGAC 4075
QY 5463 GAGCATCAAAAAATCCAGCTCAAGTCAAGTCAAGTGGCGAAACCCGACAGGACTATAAAGA 5522

Db 4076 GAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATATAAGA 4135
 Qy 5523 TACCAGGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTGCGCTT 5582
 Db 4136 TACCAGGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCCCTGTTCCGACCCCTGCGCTT 4195
 Qy 5583 ACCGATACCTGTCGCCCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC 5642
 Db 4196 ACCGATACCTGTCGCCCTTTCTCCCTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC 4255
 Qy 5643 TGTAGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCCAAAGTGGCTGTGTGACGAAACC 5702
 Db 4256 TGTAGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCCAAAGTGGCTGTGTGACGAAACC 4315
 Qy 5703 CCGTTCAGCCGACCGCTGCGCTTATCGGTAACTATCGTTCGAGTCCAAACCGGTA 5762
 Db 4316 CCGTTCAGCCGACCGCTGCGCTTATCGGTAACTATCGTTCGAGTCCAAACCGGTA 4375
 Qy 5763 AGACAGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTAT 5822
 Db 4376 AGACAGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTAT 4435
 Qy 5823 GTAGCGGTCTACAGAGTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGGACA 5882
 Db 4436 GTAGCGGTCTACAGAGTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGGACA 4495
 Qy 5883 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTCT 5942
 Db 4496 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTCT 4555
 Qy 5943 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATT 6002
 Db 4556 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATT 4615
 Qy 6003 ACGCGCAGAAAAAAGATCTCAAGAAATGCTTTTGAATCTTTTCTACGGGGCTGACGCT 6062
 Db 4616 ACGCGCAGAAAAAAGATCTCAAGAAATGCTTTTGAATCTTTTCTACGGGGCTGACGCT 4675
 Qy 6063 CAGTGAACGAAACTCAGTTAAGGATTTGGTTCATGAGATTATCAAAAGGATCTTC 6122
 Db 4676 CAGTGAACGAAACTCAGTTAAGGATTTGGTTCATGAGATTATCAAAAGGATCTTC 4735
 Qy 6123 ACTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAA 6182
 Db 4736 ACTAGATCTTTT-----GGCGCGCGCCAAATCAATCTAAAGTATATATGAGTAA 4788
 Qy 6183 ACTTGTCTGACAGTTACAAATGCTTAATCAGTAGGCACTTATCTCAGCGATCTGTCTA 6242
 Db 4789 ACTTGTCTGACAGTTACAAATGCTTAATCAGTAGGCACTTATCTCAGCGATCTGTCTA 4848
 Qy 6243 TTTCTGTTCCATAGTTGCTGACTCCCGTGGTGTAGATAACTACGATACGGAGGCG 6302
 Db 4849 TTTCTGTTCCATAGTTGCTGACTCCCGTGGTGTAGATAACTACGATACGGAGGCG 4908
 Qy 6303 TTACCATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGAT 6362
 Db 4909 TTACCATCTGGCCCGAGTGTGCAATGATACCGGAGACCCAGCTCACCGGCTCCAGAT 4968
 Qy 6363 TTATCAGCAATAAACACGACCGAGCGGAAGGCGGAGCGAGAGTGGTCTGCAACTTTA 6422
 Db 4969 TTATCAGCAATAAACACGACCGAGCGGAAGGCGGAGCGAGAGTGGTCTGCAACTTTA 5028
 Qy 6423 TCGGCTCCATCCAGTCTATTAAATGTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTT 6482
 Db 5029 TCGGCTCCATCCAGTCTATTAAATGTTGGCGGAAGCTAGAGTAAGTAGTTCGCCAGTT 5088
 Qy 6483 AATAGTTTGGCAACAGTGTGTGCAATGCTTACAGGC-TGCTGTGTGTCAGCTCGTGGTT 6541
 Db 5089 AATAGTTTGGCAACAGTGTGTGCAATGCTTACAGGCATCGTGTGTGTCAGCTCGTGGTT 5148
 Qy 6542 GGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATG 6601
 Db 5149 GGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGGAGTTACATGATCCCCCATG 5208

Qy 6602 TTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGTCAGAGTAAGTTGGCC 6661
 Db 5209 TTGTGCAAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGTCAGAGTAAGTTGGCC 5268
 Qy 6662 CGAGTGTATCACTCATGTTATGCGACACTGCAATAATCTCTTACTGTCTCATGCAATCC 6721
 Db 5269 CGAGTGTATCACTCATGTTATGCGACACTGCAATAATCTCTTACTGTCTCATGCAATCC 5328
 Qy 6722 GTAAAGATGCTTTTCTGTGACTGGTGAAGTCTCAACCAAGTCAATCTGAGAAATAGTATG 6781
 Db 5329 GTAAAGATGCTTTTCTGTGACTGGTGAAGTCTCAACCAAGTCAATCTGAGAAATAGTATG 5388
 Qy 6782 CGCGACCGAGTGTCTTTGCGCGCGCTCAATACCGGATAATACCGCGGCACATAGCAGA 6841
 Db 5389 CGCGACCGAGTGTCTTTGCGCGCGCTCAATACCGGATAATACCGCGGCACATAGCAGA 5448
 Qy 6842 ACTTTAAAGTCTCATCATTTGGAAGAGTCTTCTCGGGCGGAAAACTCTCAAGGATCTTTA 6901
 Db 5449 ACTTTAAAGTCTCATCATTTGGAAGAGTCTTCTCGGGCGGAAAACTCTCAAGGATCTTTA 5508
 Qy 6902 CGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCACCCCACTGATCTTCAGCATCT 5961
 Db 5509 CGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCACCCCACTGATCTTCAGCATCT 5568
 Qy 6962 TTTTACTTTTCAACAGCGTCTTCTGGGTGAGCAAAAGAGGCAAAATGCGGCAAAAAAG 7021
 Db 5569 TTTTACTTTTCAACAGCGTCTTCTGGGTGAGCAAAAGAGGCAAAATGCGGCAAAAAAG 5628
 Qy 7022 GGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATTATGA 7081
 Db 5629 GGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATTATGA 5688
 Qy 7082 AGCATTTATCAGGTTATGCTCTCATGAGCGATACATATTTGAATGATTTAGAAAAAT 7141
 Db 5689 AGCATTTATCAGGTTATGCTCTCATGAGCGATACATATTTGAATGATTTAGAAAAAT 5748
 Qy 7142 AAACAAATAGGGGTTCCGCGCACATTTCC 7170
 Db 5749 AAACAAATAGGGGTTCCGCGCACATTTCC 5777

RESULT 7
 E23357
 LOCUS 7372 bp DNA linear PAT 18-JUN-2001
 DEFINITION Virus vector system expressing apoptosis-related gene.
 ACCESSION E23357
 VERSION E23357.1 GI:13024380
 KEYWORDS JP 1999075859-A/3.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 7372)
 AUTHORS Hirofumi,H.
 TITLE Virus vector system expressing apoptosis-related gene
 JOURNAL Patent: JP 1999075859-A 3 23-MAR-1999;
 R B R JENSEN KK
 COMMENT OS Homo sapiens (human)
 PN JP 1999075859-A/3
 PD 23-MAR-1999
 PF 08-SEP-1997 JP 1997259235
 PR
 PI HIROFUMI HAMADA
 PC C12N15/09,C12N5/10,C12N7/00//A61K35/76,A61K48/00,(C12N5/10,PC
 C12R1:91),
 PC (C12N7/00,C12R1:92),C12N15/00,C12N5/00,(C12N5/00,C12R1:91) CC
 Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..7372 /organism='Homo sapiens (human)'.
 FT Location/Qualifiers

source	1. .7372	Query Match	41.6%;	Score 3183.2;	DB 6;	Length 7372;
		Best Local Similarity	70.7%;	Pred. No. 0;		
		Matches 4893;	Conservative 0;	Mismatches 1503;	Indels 528;	Gaps 28;
ORIGIN						
Qy	843	TTCTAGAGAACCATCAGATGTTTCCAGGGTCCGCCAAGGACCTCAAAATGACCTGTGCCT	902			
Db	467	TGCTAGAGAACCATCAGATGTTTCCAGGGTCCGCCAAGGACCTCAAAATGACCTGTGCCT	526			
Qy	903	TATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTCCCGGAG	962			
Db	527	TATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTCCCGGAG	586			
Qy	963	CTCAATAAAGAGCCCAACCCCTCACTGGCGCGCCAGTCTCTCCGATGACGTGCGTCG	1022			
Db	587	CTCAATAAAGAGCCCAACCCCTCACTGGCGCGCCAGTCTCTCCGATGACGTGCGTCG	646			
Qy	1023	CCGGGTACCGGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGGACCTGC	1082			
Db	647	CCGGGTACCGGTATTTCCCAATAAAGCCTCTTGTGTTTGCATCCGAATCGTGGTCCTGC	705			
Qy	1083	TGATCTTTGGAGGGTCTCTTCAGATTGATGACTGCCC-ACCTCGGGGGTCTTTCAATTT	1141			
Db	706	TGTTCTTTGGAGGGTCTCTCTGAGTGTGACTACCCGTCAGCGGGGTCTTTCAATTT	765			
Qy	1142	GGAGGTTCCACCGAGATTGAGAACCCCTGCCCCAGGACCAACGACCCCCCGCCGGAG	1201			
Db	766	GGGGGCTCGTCCGGGATCGGAGACCCCTGCCCCAGGACCAACCGA-CCACACACCGGGAG	824			
Qy	1202	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCGGAATGTCTAGTGTCTATGACTGATTTT	1261			
Db	825	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCGGAATGTCTAGTGTCTATGACTGATTTT	884			
Qy	1262	ATGCGCTCGGTCCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGGA	1321			
Db	885	ATGCGCTCGGTCCGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGGA	944			
Qy	1322	ACTGAACGATTCGGAACACCCGGCCGCAACCCCTGGGAGACGTCACGGGACTTCGGGGGC	1381			
Db	945	ACTGACGATTCGGAACACCCGGCCGCAACCCCTGGGAGACGTCCTCCAGGACTTCGGGGGC	1004			
Qy	1382	CGTTTTTGTGCCCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGTTGTCACC	1441			
Db	1005	CGTTTTTGTGCCCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGTTGTCACC	1064			
Qy	1442	CCCCTTAGAGAGGGGATATGGTTCCTGTTAGGAGACGAGAACCTAAACAGTTTCCCGCC	1501			
Db	1065	CCCCTTAGAGAGGGGATATGGTTCCTGTTAGGAGACGAGAACCTAAACAGTTTCCCGCC	1124			
Qy	1502	TCCGCTCTGAATTTTTCGTTTTCGGTTTGGGACCCGAAGCCGCGCGCGGCTCTTGTCTGCTG	1561			
Db	1125	TCCGCTCTGAATTTTTCGTTTTCGGTTTGGGACCCGAAGCCGCGCGCGGCTCTTGTCTGCTG	1184			
Qy	1562	CAGCATCGTTCTGTGTTGTCTGTCTGTGACTGTGTTTGTGTTTGTCTGAAATATGGG	1621			
Db	1185	CAGCATCGTTCTGTGTTGTCTGTCTGTGACTGTGTTTGTGTTTGTCTGAAATATAGGG	1244			
Qy	1622	CCGGGGCTAGACTGTATACCTCCCTTAAGTTTTCACCTTAGGTCACCTGGAAGATGTCGA	1681			
Db	1245	CC-----AGACTGTATACCTCCCTTAAGTTTTCACCTTAGGTCACCTGGAAGATGTCGA	1298			
Qy	1682	GCGGATCGCTCACACCAAGTCGGTGTGATGTCTGATGTTTGTGTTTGTCTGAAATATGGG	1741			
Db	1299	GCGGATCGCTCACACCAAGTCGGTGTGATGTCTGATGTTTGTGTTTGTCTGAAATATGGG	1358			
Qy	1742	TGCAAAATGCGCAACCTTTTAACTGTCGATGCGCGCGAGACCGCACTTTTAAACCGAGACCT	1801			
Db	1359	TGCAAAATGCGCAACCTTTTAACTGTCGATGCGCGCGAGACCGCACTTTTAAACCGAGACCT	1418			

Db	4649	AA	4700
Qy	4675	GCCACACGACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCCT	4734
Db	4709	AA	4768
Qy	4735	CGTCTATTTTGTGGCTTCTCCGTATTTTAAATTTCTAGTTTGCATCCCTTCTCTGAG	4794
Db	4769	AA	4828
Qy	4795	CACGCGATTGACGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAGGTTGCCTGG	4854
Db	4829	AA	4888
Qy	4855	GCTCAGTGTGAGATTTTGGCCATAAAAGGGGTCTCTGCCCTGTGTACAGACAGATCGAA	4914
Db	4889	AA	4948
Qy	4915	TCTAGATGTCATCTCAGAGTCCCGGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGC	4974
Db	4949	AA	5008
Qy	4975	CTAGAGATCTCTACGCGGACGATCGTGGCCGGGTACCGAGCTCGAATTCGTAATCAT	5034
Db	5009	AA	5068
Qy	5035	GGTCATAGCTGTTTCTGTGTGAAATTTGTATCCGCTCACAATTTCCACACAATACGAG	5094
Db	5069	AA	5128
Qy	5095	CCGGAAGCATAAAGTGTA-----AG	5115
Db	5129	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGTCGAAGCTTGGCGAGATTTTCAGGAG	5188
Qy	5116	CCTGGGTGCTATAGTAGTGAGCTAACTCACAATTAATTTGGTTGGCTCACTGCC-----	5171
Db	5189	CTAAGGAAGCTAAATGGAGAAAAAATCATCTGGATATCACCGTTTGATATATCCCAAT	5248
Qy	5172	-----GTTTTCCAGTCGGGAAACCTGT	5193
Db	5249	GGCATCGTAAGAACAATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTACTATAACAGA	5308
Qy	5194	CGTGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTTCGTATTTGGGC	5253
Db	5309	CCGTTACGTGCAATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTTCGTATTTGGGC	5368
Qy	5254	GCTCTTCGGTCTCTCGCTCACTGACTCGCTCGGCTCGGTCGGCTCGGCGGAGCGG	5313
Db	5369	GCTCTTCGGTCTCTCGCTCACTGACTCGCTCGGCTCGGTCGGCTCGGCGGAGCGG	5428
Qy	5314	TATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAA	5373
Db	5429	TATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAAGAA	5488
Qy	5374	AGAAATGTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGCTTCTCTGG	5433
Db	5489	AGAAATGTGAGCAAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAGCCCGCTTCTCTGG	5548
Qy	5434	CGTTTTTCCATAGGCTCGGCCCTCGACAGAGATCAAAAAATCGACGCTCAAGTCAGA	5493
Db	5549	CGTTTTTCCATAGGCTCGGCCCTCGACAGAGATCAAAAAATCGACGCTCAAGTCAGA	5608
Qy	5494	GGTGGCAAAACCGACAGGACTATAAGATACAGAGGTTTTCCCTCGGAGCTCCCTCG	5553
Db	5609	GGTGGCAAAACCGACAGGACTATAAGATACAGAGGTTTTCCCTCGGAGCTCCCTCG	5668
Qy	5554	TGCGCTCTCTGTTCCGACCTCGCGTTTACCGGATACCTGTCCGCTTTTCTCCCTTCGG	5613
Db	5669	TGCGCTCTCTGTTCCGACCTCGCGTTTACCGGATACCTGTCCGCTTTTCTCCCTTCGG	5728
Qy	5614	GAACGTGGCGCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTCCGTTAGGTCGTTTC	5673
Db	5729	GAACGTGGCGCTTTCTCAATGTCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTCGTTTC	5788

Qy	5674	GCTCAAGCTGGGTGTGTGCAAGAAACCCCGCTTACAGCCGACCGCTGCGCCTTATCCG	5733
Db	5789	GCTCAAGCTGGGCTGTGTGCACGAAACCCCGCTTACAGCCGACCGCTGCGCCTTATCCG	5848
Qy	5734	GTAACATCGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA	5793
Db	5849	GTAACATCGTCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA	5908
Qy	5794	CTGGTAAACAGGATTAGCAGACGCGAGGTATGTAGCGCGTGTCTACAGAGTCTCTGAAGTGGT	5853
Db	5909	CTGGTAAACAGGATTAGCAGACGCGAGGTATGTAGCGCGTGTCTACAGAGTCTCTGAAGTGGT	5968
Qy	5854	GGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAG	5913
Db	5969	GGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAG	6028
Qy	5914	TTACCTTTCCGAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCGCTGGTAGCG	5973
Db	6029	TTACCTTTCCGAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCGCTGGTAGCG	6088
Qy	5974	GTGGTTTTTTTGTGTGCAAGCAGCAGATTAACGGCGCAGAAAAAAGGATCTCAAGAAGATC	6033
Db	6089	GTGGTTTTTTTGTGTGCAAGCAGCAGATTAACGGCGCAGAAAAAAGGATCTCAAGAAGATC	6148
Qy	6034	CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAAGGGATT	6093
Db	6149	CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAAACTCACGTTAAAGGGATT	6208
Qy	6094	TGGTCATGAGATTATCAAAAAGGATCTTCACTTAGATCTCTTTTAAATTAATAATGAAGTT	6153
Db	6209	TGGTCATGAGATTATCAAAAAGGATCTTCACTTAGATCTCTTTTAAATTAATAATGAAGTT	6268
Qy	6154	TTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCA	6213
Db	6269	TTAAATCAATCTAAAGTATATATAGTAAACTTTGGTCTGACAGTTACCAATGCTTAATCA	6328
Qy	6214	GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTGACTCCCCG	6273
Db	6329	GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCATAGTTGCTGACTCCCCG	6388
Qy	6274	TCGTGTAGATAACTACGATHACGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATAC	6333
Db	6389	TCGTGTAGATAACTACGATHACGGAGGGCTTACCATCTGGCCCCAGTGTCTGCAATGATAC	6448
Qy	6334	CGGAGACCCACGCTCACCGGCTCCAGATTATACGCAATTAACACGACGACGCGGAGGG	6393
Db	6449	CGGAGACCCACGCTCACCGGCTCCAGATTATACGCAATTAACACGACGACGCGGAGGG	6508
Qy	6394	CCGAGCGCAAGTGGTCTGCAACTTATATCCGCCCTCCATCCAGTCTATTAATTTGTGCC	6453
Db	6509	CCGAGCGCAGCGTGGTCTGCAACTTATATCCGCCCTCCATCCAGTCTATTAATTTGTGCC	6568
Qy	6454	GGGAAGCTAGATAAGTAGTTTGGCCAGTTTAATAGTTTGGCAACGTTTGTGCCATTTGCTA	6513
Db	6569	GGGAAGCTAGATAAGTAGTTTGGCCAGTTTAATAGTTTGGCAACGTTTGTGCCATTTGCTA	6628
Qy	6514	CAGGC-TCGGTGGTCAACGCTCGTCTGTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAC	6572
Db	6629	CAGGCATCGTGGTCAACGCTCGTCTGTTGGTATGGCTTCATTCAGTCTCGGTTCCCAAC	6688
Qy	6573	GATCAAGGGCAGTATACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTCGGTC	6632
Db	6689	GATCAAGGGCAGTATACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCTCGGTC	6748
Qy	6633	CTCCGATCGTTGTCAAGAAGTAAGTTGGCCCGCAGTGTATCACTCAATGGTTATGGCAGCAC	6692
Db	6749	CTCCGATCGTTGTCAAGAAGTAAGTTGGCCCGCAGTGTATCACTCAATGGTTATGGCAGCAC	6808
Qy	6693	TGCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACT	6752
Db	6809	TGCATAAATCTCTTACTGTCAATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAAGTACT	6868

6753 CAACCAAGTCATCTGAGAAATAGTGTATGCGGACCGAGTGTCTCTTCCCGCGGTCAA 6812
 Db CAACCAAGTCATCTGAGAAATAGTGTATGCGGACCGAGTGTCTCTTCCCGCGGTCAA 6928
 6813 TAGGGATTAATCCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAACGTT 6872
 Db TAGGGATTAATCCGCGGCACATAGCAGAACTTTAAAGTGTCTCATCATTTGAAACGTT 6988
 6873 CTTCCGGGGGAAACCTCTCAAGGATCTTACCGGTGTGAGATCCAGTTCGATGTAACCCA 6932
 Db CTTCCGGGGGAAACCTCTCAAGGATCTTACCGGTGTGAGATCCAGTTCGATGTAACCCA 7048
 6933 CTCGTCCACCAACTGATCTTTCAGCATCTTTTACTTTTCCACGCGTTTCTCGGTGAGCAA 6992
 Db CTCGTCCACCAACTGATCTTTCAGCATCTTTTACTTTTCCACGCGTTTCTCGGTGAGCAA 7108
 6993 AAACAGGAAGGCAAAATGCGCAAAAAGGAATAAGGCGGACACGCAAAATGTTGAATAC 7052
 Db AAACAGGAAGGCAAAATGCGCAAAAAGGAATAAGGCGGACACGCAAAATGTTGAATAC 7168
 7053 TCATCTCTCTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTTCTCATGAGCG 7112
 Db TCATCTCTCTCTTTTCAATATTTATGAAGCATTTATCAGGGTTATTTCTCATGAGCG 7228
 7113 GATACATATTGAAATGATTTAGAAAATAAACAATAGGGTTCCGCGCACATTTCCCC 7172
 Db GATACATATTGAAATGATTTAGAAAATAAACAATAGGGTTCCGCGCACATTTCCCC 7288
 7173 GAAAGTGCACCTGACGTCTAAGAAACCATTTATATCATGATTAACCTATAAAATA 7232
 Db GAAAGTGCACCTGACGTCTAAGAAACCATTTATATCATGATTAACCTATAAAATA 7348
 7233 GCGGTATCAGGCGGCTTTCGTC 7256
 Db GCGGTATCAGGCGGCTTTCGTC 7372

RESULT 8
 AB086385
 LOCUS AB086385 6277 bp DNA circular SYN 14-NOV-2003
 DEFINITION Retroviral vector pCX4neo DNA, complete sequence.
 ACCESSION AB086385
 VERSION AB086385.1 GI:26449191
 KEYWORDS Retroviral vector pCX4neo
 SOURCE Retroviral vector pCX4neo
 ORGANISM other sequences; vectors.
 REFERENCE 1
 Akagi, T., Sasai, K. and Hanafusa, H.
 Refractory nature of normal human diploid fibroblasts with respect
 to oncogene-mediated transformation
 Proc. Natl. Acad. Sci. U.S.A. 100 (23), 13567-13572 (2003)
 MEDLINE 22975258
 PUBMED 14597713
 REFERENCE 2 (bases 1 to 6277)
 Akagi, T.
 Direct Submission
 Submitted (10-JUN-2002) Tsuyoshi Akagi, Osaka Bioscience Institute,
 Molecular Oncology; Furuedai 6-2-4, Suita, Osaka 565-0874, Japan
 (E-mail: takagi@obi.or.jp, Tel: 81-6-6872-4834, Fax: 81-6-6871-7521)
 Location/Qualifiers
 1. 6277
 /organism="Retroviral vector pCX4neo"
 /mol_type="other DNA"
 /db_xref="taxon:197851"
 1. 466
 /notes="CMV IE enhancer"
 467. 762
 /note="R-US region of MuLV 5' LTR"
 763. 1667
 /note="psi+; extended packaging signal"
 1668. 2045
 /note="Mo-MuLV pol region containing splice acceptor"

misc_feature 2044. 2106
 /note="multiple cloning site"
 RBS 2107. 2716
 /note="IRES sequence from EMCV"
 gene 2718. 3512
 /gene="neo"
 CDS 2718. 3512
 /gene="neo"
 /codon_start=1
 /product="aminoglycoside phosphotransferase"
 /protein_id="BAC41724.1"
 /db_xref="GI:26449192"
 /translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGSDAAVFLRLSAQGR
 PVLFKTDLSGALNELODEAARLSWLTTPVCAAVLDVVTAGRDWLLGLGVPQDL
 LSHLAPEAKYSIMADAMRGLHTLOPATCPDPAKHRIERARTMEAGLVDDDDLE
 EHOGLAPAEFLFARLKARPDGEDLVTHGGDACLPNIMVENGFRSFGIDCGRGLGVADRY
 QDIALATRDIAELGGEWADRLVLYGIAAPDSQRIAFYRLLEDF"
 3572. 4163
 /note="3' LTR of Mo-MuLV"
 complement(5216. 6076)
 gene /gene="bla"
 complement(5216. 6076)
 CDS /gene="bla"
 /codon_start=1
 /product="beta-lactamase"
 /protein_id="BAC41725.1"
 /db_xref="GI:26449193"
 /translation="MSIOHFRVALIPFFAFCFLPVFAHPETLVKVKDAEDQLGARVGY
 IELDINSKGLLSEFRPEFPMSTFKVLLGAVLSRVDDAQEGQELRRIRHYSQNDLVE
 YSPVTEKLTDMGTIVRELCSAAITMSDNTAANLLTTGGPKELTAFLHNMDDHVTFL
 DRWPELNEATIPNDRDTPVAMATTLKLTGELLTLASQQILDMWEADKVGPL
 LRSALPAGWFIADKSGAGERSGIIAALPGDKPSRIWVIYTTGSAQTMDERNRQIA
 EIGASLIKHW"
 ORIGIN
 Query Match 41.5%; Score 3179.2; DB 12; Length 6277;
 Best Local Similarity 73.1%; Pred. No. 0;
 Matches 4719; Conservative 0; Mismatches 1048; Indels 691; Gaps 22;
 QY 843 TTCTAGAGAACCATCAGATGTTTCAGGGTGCCTCCAGACCTGAAATGACCTGTGCCT 902
 Db 466 TGCTAGAGAACCATCAGATGTTTCAGGGTGCCTCCAGACCTGAAATGACCTGTGCCT 525
 QY 903 TATTTGAACCTAACCAATCAGTTCTCGCTTCTGTCGCGCGCTTCTGCTCCCGGAG 962
 Db 526 TATTTGAACCTAACCAATCAGTTCTCGCTTCTGTCGCGCGCTTCTGCTCCCGGAG 585
 QY 963 CTCAATAAAGAGCCCAACCCCTCCTCCTCGGCGGCCAGTCTCCGATAGACTGCTCG 1022
 Db 586 CTCAATAAAGAGCCCAACCCCTCCTCCTCGGCGGCCAGTCTCCGATAGACTGCTCG 645
 QY 1023 CCGGGTACCGGTATTCCTCAATAAAGCCTCTGCTGTTTGGCATCGAATCGTGGACTCG 1082
 Db 646 CCGGGTACCGGTATTCCTCAATAAAGCCTCTCTTGCAG-ATTGCATCCGACTTGTGGTCTCG 704
 QY 1083 TGATCTTGGGAGGGTCTCTCAGATTGATTGACTGCCC-ACCTCGGGGTCTTTTCATTT 1141
 Db 705 TGTTCCTTGGGAGGGTCTCTCTGAGTGAATGACTACCCGTGACGGGGGTCTTTTCATTT 764
 QY 1142 GGAGGTTTCCACGAGATTGGAGACCCCTCGCCAGGAGCCACCGACCCCGCCCGGGAG 1201
 Db 765 GGGGGCTGCTCCGGGATCGGAGACCCCTCGCCAGGAGCCACCGA-CCCAACACCGGGAG 823
 QY 1202 GTAAGCTGGCCAGCAACTTATCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
 Db 824 GTAAGCTGGCCAGCAACTTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883
 QY 1262 ATGCGCTGCGTCGCTACTAGTTAGTAACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
 Db 884 ATGCGCTGCGTCGCTACTAGTTAGTAACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
 QY 1322 ACTGACGAGTTCCGAAACACCCCGCGCAACCTCGGAGACGCTCCAGGAGACTTCGGGGGC 1381

Db 944 ACTGACGAGTTCCGGAACACCCCGCGCAACCTCGGAGACGTCCTCAGGAGCTTCGGGGGC 1003
Qy 1382 GCTTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGTGTGACCC 1441
Db 1004 GCTTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGTGTGACCC 1063
Qy 1442 CCCCTAGAGAGGATATGTTCTCTGTTAGGAGACGAGAACCTTAAACAGTTTCCCGCC 1501
Db 1064 CCCCTAATAGAGGATATGTTCTCTGTTAGGAGACGAGAACCTTAAACAGTTTCCCGCC 1123
Qy 1502 TCCGCTGTAATTTTGTCTTCCGTTTGGAGCCGAGCCGCGCGCTCTTGTCTGCTG 1561
Db 1124 TCCGCTGTAATTTTGTCTTCCGTTTGGAGCCGAGCCGCGCGCTCTTGTCTGCTG 1183
Qy 1562 CAGCATCGTCTGTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621
Db 1184 CAGCATCGTCTGTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243
Qy 1622 CCCGGCTAGACTGTTACCACTCCCTTAAGTTTACCTTAGGTCACTGGAAGATGCGA 1681
Db 1244 CC-----AGACTGTTACCACTCCCTTAAGTTTACCTTAGGTCACTGGAAGATGCGA 1297
Qy 1682 GCGGATCGCTCACAACCACTCGGTAGATGTTCAAGAGAGACGTTGGGTTACCTTCTGCTC 1741
Db 1298 GCGGATCGCTCACAACCACTCGGTAGATGTTCAAGAGAGACGTTGGGTTACCTTCTGCTC 1357
Qy 1742 TGCAGATGGCCAACTTTTAACTGGATGCGCGAGACGGCACTTTTAAACCGAGACCT 1801
Db 1358 TGCAGATGGCCAACTTTTAACTGGATGCGCGAGACGGCACTTTTAAACCGAGACCT 1417
Qy 1802 CATACCCAGGTTAAGATCAAGGCTTTTCACTTGGCCCGCATGAGACCCAGACGAGGT 1861
Db 1418 CATACCCAGGTTAAGATCAAGGCTTTTCACTTGGCCCGCATGAGACCCAGACGAGGT 1477
Qy 1862 CCCCTACATCGTACCTCGGAGGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGACCTT 1921
Db 1478 CCCCTACATCGTACCTCGGAGGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGACCTT 1537
Qy 1922 TGTAACCCCTAAGCCTCGCCTCTCTTCCATCGCCCGCCGCTCTCTCCCGCTTGAACC 1981
Db 1538 TGTAACCCCTAAGCCTCGCCTCTCTTCCATCGCCCGCCGCTCTCTCCCGCTTGAACC 1597
Qy 1982 TCCTCGTTGACACCCGCTCGATCGTCTCTTATTCAGCCCTCACTCTCTCTAGGGCC 2041
Db 1598 TCCTCGTTGACACCCGCTCGATCGTCTCTTATTCAGCCCTCACTCTCTCTAGGGCC 1657
Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101
Db 1658 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 1717
Qy 2102 CCCTGACATGACAAGAGTTACTAAGAGCCCTCTCTCAAGCTCACTTACAGGCT-TCTA 2160
Db 1718 CCCTGACTTGACAAGAGTTACTAAGAGCCCTCTCTCAAGCTCACTTACAGGCTCTCTA 1777
Qy 2161 CTTAGTCAGCAGAGGCTTGAGACCTCTGGGGCAGCCTTCAAGAACAACTTGGACCG 2220
Db 1778 CTTAGTCAGCAGAGGCTTGAGACCTCTGGGGCAGCCTTCAAGAACAACTTGGACCG 1837
Qy 2221 ACCGGTGGTCACTCACTTACAGTCCGGGACACAGTGTGGGTCCGCGACACAGAC 2280
Db 1838 ACCGGTGGTCACTCACTTACAGTCCGGGACACAGTGTGGGTCCGCGACACAGAC 1897
Qy 2281 TAAGAACCTAGAACCTCTGCTGGAAGGACCTTACACAGTCTCTGTGACCACTCCCGACCGC 2340
Db 1898 TAAGAACCTAGAACCTCTGCTGGAAGGACCTTACACAGTCTCTGTGACCACTCCCGACCGC 1957
Qy 2341 CCTCAAAGTAGACGGCATCGAGTTGGATACACCGCCGCCACGTGAAGGCTGCCGACCC 2400
Db 1958 CCTCAAAGTAGACGGCATCGAGTTGGATACACCGCCGCCACGTGAAGGCTGCCGACCC 2017
Qy 2401 CGGGGTGGACCATCTCTAGACTGCC-----ATGGGATGAGCTGTATCATCTCT 2452
Db 2018 CGGGGTGGACCATCTCTAGACTGCCCGGATCCAGTGTGGGTGAGCGGAACTTCCGGAC 2077

Qy 2453 TCCTTGTAGCAACAGCTACAGGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGCA 2512
Db 2078 CGACCGCTGGCGCGCGGTAAACAATTGTTAACTAACTTAAGCTAGCAACGGTTTCCCTCT 2137
Qy 2513 GCCTTAGCGCCAGCGTGGGTGACAGAGTGACCATCACCTGTAAAGCCAGTCAGGATGTGG 2572
Db 2138 AGCGGATCAATTCGCGCCCCCCCCCTTAACGTTACTTGGCCGAAGCCGCTTGGAAATAAGG 2197
Qy 2573 GTACTTCTGTAGCTTGGTACCCAGCAGAACCCAGGTCTCAAAAGCTGCTGATCTACT 2632
Db 2198 CCGGTGTGCGTTTGTCTATATGTTTATTTTCCACCATATGCGCGTCTTTTGGCAATGTGAG 2257
Qy 2633 GGACATCAACCCGCAACACTGCTGTGCCAAGCAGATTACAGCGGTAGCGGTACCG 2692
Db 2258 GGCCCGGAAACCTGGCCCTGTCTTCTTGAAGCAGCATCTCTAGGGGTCTTTTCCCTCTCGC 2317
Qy 2693 ACTTCACTTCACTCAGCAGCCCTCCAGCCAGAGGACATCGCCACTACTACTTGCACG 2752
Db 2318 CAAAGGAATGCAAGGTCTGTTGAATGCTGTAAGAAAGAGTTCCTCTTGGAAAGCTTCTTG 2377
Qy 2753 AATATAGCCTCTATTCGGTCG-----TTCCGCCAAGGGACCAAGGTGGAAATCAAAACG 2804
Db 2378 AAGACAACAACTCTGTAGCGACCTTTGCAGCAGCGGAAACCCCCACCTCTGGGACAG 2437
Qy 2805 AGTGGCTCAGGATCGGGTGGATCCGGCTCTGTGGTCTAGGATCGGAGGTCCAACTGGT 2864
Db 2438 GTGCTCTGGCGCCAAAGGCCAGCTGTATAAGATACACCTGCAAAAGCGGCACAAACCCA 2497
Qy 2865 GGAGAGCGGTGGAGTGTGTGCAACCTGGCCGCTCTCTGGCCCTGCTCTCTCGGATC 2924
Db 2498 GTGCCAGTGTGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCTCAAGCGTATT 2557
Qy 2925 TGGCTTCGATTTACCACATATTGGAATGAGTTGGGTGAGACAGGCACCTGCAAAAGGTCT 2984
Db 2558 CAACAAGGGCTGAGGATGCCAGAAAGGTACCCCATTTGATGGAATCTGATCTGGGGCC 2617
Qy 2985 TGAAGTGGATGGAAATTCATCCAGATAGCAGTACGATTAACTATATGCGCGTCTCTAAA 3044
Db 2618 TCGGTGCACATGCTTTACGTGTGTTAGTCAGGTTTAAATAACGCTCTAGGCCCCCGAAC 2677
Qy 3045 GGATAGATTTACATATATCGGAGACACGCCAAGAACACATTTCTCTGCA-- 3096
Db 2678 CAGCGGACGTGGTCTTCTTTTGAATAACAGATACCATGATGAACAAAGATGGATT 2737
Qy 3097 --ATGGACAGCTCAGACCCGAGACACCGGGTCTATTTTGTGCAAGCTTTACTTCG 3154
Db 2738 GCACGAGGTTCTCCGCGCGCTTGGGTGGAGAGCTATTGGCTATGACTGGGCACACA 2797
Qy 3155 G--CTTCCCTGGTCTTATTTGGGCCAAGGGACCCCGTCAACCGTCTCCAGTGTAA 3212
Db 2798 GACAATCGGCTGCTCTGATGCGCGGTGTTCCGCTGTGACGCGAGGGGCCCGGTTCT 2857
Qy 3213 GCCCACAAGACCGCAGCGCGGACCAACCAACCGCGGCCACCATCGCGTGCAGCC 3272
Db 2858 TTTTGTCAAGACACCTGTCCGCTGCTGCTGAACTGACGAGCAGGCGAGCGGCT 2917
Qy 3273 CTTGCTCCCTCGCCAGAGCGGCTCGGCCAGCGCGGGCGGCGAGTGCACAC--GAG 3329
Db 2918 ATGTTGCTGGCCAGACGGCGGTTCTTGGCGAGCTGTGCTCGAGTTGCTACTGAAGC 2977
Qy 3330 GGGGCTGGAATTCGCGCTTGGATCCCAAACTCTGTACTCTGTGATGGAATCTCTTCAT 3389
Db 2978 GGGAGGACTGCGTCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCT 3037
Qy 3390 CTATGTTGTCATTCTCATCTGCTTGTCTGAGAGTGAAGTTGACGAGGAGCGGACGCC 3449
Db 3038 TGCTCTCGCGAGAAAGTATCCATATCGGTGATGCAATGCGCGGCTGCTATACGCTTGA 3097
Qy 3450 CCGCGGTACAGCAGGCGGCGAGACCACTTAAACGAGCTCAATCTTAGGACCAAGAGA 3509
Db 3098 TCCGGCTACTTGCCTTCCAGCCACCAAGGAAACATCGCATCGAGGAGCAGCTACTCG 3157

QY 3510 GGAGTACGATGTTTGGACAAGAGACGTGGCCGGGACCTGAGATGGGGGAAAGCCGAG 3569
 Db 3158 GATGGAAGCGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGG--GCTCGCG 3215
 QY 3570 AAGGAAGAACCTCTAGGAAGGCTGTACAATGAATGCAAGAGATAAGATGGCGGAGGC 3629
 Db 3216 CCAGCGCAACTGTCTGCCAGGCTCAAGGGCGCATGCCGAGCGGAGGATCTCGTCGTG 3275
 QY 3630 CTACAGTGAATGGATGAAGGCGAGCGCGGAGGGGCAAGG--GGCAGATGGCCTT 3687
 Db 3276 ACCATGGGAGCTCTGCTTGGCAATATCATGGTGAATAATGGCGCTTTCTGGATTC 3335
 QY 3688 TACCAGGCTCTAGTACAGCACCAAGGACACTACGACGCCCTTCACATGACGGGCCCTG 3747
 Db 3336 ATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCGTTGCTACCCGT 3395
 QY 3748 CCCCCCTCGCTAA---CTCGACGCGCGCGGATCCGATTAAGTCCAAATTTGTTAAAGAC 3803
 Db 3396 GATATTGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCTGTGCTTACGGTATC 3455
 QY 3804 AGGATATCAGTGGTCCAGGCTCTAGTTTGTGACTCAACAATATCACCAGCTGAAGCCTATA 3863
 Db 3456 GCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGAGCG 3515
 QY 3864 GAGTACGAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATG 3923
 Db 3516 GGACTCTGGGTTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATG 3575
 QY 3924 AAAGACCCACCTGTAGTTTGGCAAGCTAGTTAAGTAAACGCCAATTTTGAAGGCAATG 3983
 Db 3576 AAAGACCCACCTGTAGTTTGGCAAGCTAGTTAAGTAAACGCCAATTTTGAAGGCAATG 3635
 QY 3984 -AAATACATACTGAGATAGAGAATTCAGATCAAGTTAGGAACACA-GAGACAGCA 4041
 Db 3636 AAAAATACATACTGAGATAGAGAATTCAGATCAAGTTAGGAACACAAGTGAACAGCT 3695
 QY 4042 GAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCC--GCTCAGGGCCAAGA 4100
 Db 3696 GAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCCGCTCAGGGCCAAGA 3755
 QY 4101 ACAGTTTGAACAGGAGATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCC 4160
 Db 3756 ACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCC 3815
 QY 4161 GGCTCAGGCGCAAGACAGATGGTCCCGAGATGCGGTCCCGCCTCAGCAGTTCTAGAG 4220
 Db 3816 GGCTCAGGCGCAAGACAGATGGTCCCGAGATGCGGTCCCGCCTCAGCAGTTCTAGAG 3875
 QY 4221 AACCATCAGATGTTTCCAGGTTGCCCAAGGACCTGAAATGACCTGTGCTTATTGAA 4280
 Db 3876 AACCATCAGATGTTTCCAGGTTGCCCAAGGACCTGAAATGACCTGTGCTTATTGAA 3935
 QY 4281 CTAAACCAATCAGTTCGCTTCTCGCTTCTGTGCGCGCTTCTGCTCCCGAGCTCAATAA 4340
 Db 3936 CTAAACCAATCAGTTCGCTTCTCGCTTCTGTGCGCGCTTCTGCTCCCGAGCTCAATAA 3995
 QY 4341 AAGAGCCCAACACCCCTCAGTGGCGGCGCAGTCTCCGATAGACGTGGTCCGCGGTA 4400
 Db 3996 AAGAGCCCAACACCCCTCAGTGGCGGCGCAGTCTCCGATAGACGTGGTCCGCGGTA 4055
 QY 4401 CCGGTGTTCTCAATAAACCTCTTGAGTTGATCCGACTCGTGGTCTCGCTGTTCTTG 4460
 Db 4056 CCGGTGTTCTCAATAAACCTCTTGAGTTGATCCGACTCGTGGTCTCGCTGTTCTTG 4115
 QY 4461 GGAGGCTCT-CTCTGAGTGAATGACTACCCGTCAGCGGGTCTTTTCAGTTTCTCCACCT 4519
 Db 4116 GGAGGCTCTCTCTGAGTGAATGACTACCCGTCAGCGG----- 4155
 QY 4520 ACAGGCTCTCAATAACTCTGATGTGCGCGCAGGGAATCCGCTCAGCGGGTTTTTGT 4579
 Db 4156 ----- 4155
 QY 4580 TATAATAAATGCAAGAACAGTGTCTCCCTTCAAGCCAGACTACATCTCGACTCTCGCTT 4639

Db 4156 ----- 4155
 QY 4640 TATAAAGAAATGTTGAAGGCTCTGTGACTATCTGCCACACGACTTTTAAAGATTTTA 4699
 Db 4156 ----- 4155
 QY 4700 TGCCTCTGGATGAGGATTTAGTCAATCTATCTCGTCTATTTTGTGCTTCTCCGTA 4759
 Db 4156 -GCTTTTACATGACGATGATCAAAATTAATTTGGTTTTTTTT- 4199
 QY 4760 TTTTAAATTTCTAGTTTGCACCTCTCTGAGACACGGCGATGTCAGAGTAGTAAATA 4819
 Db 4200 ----- 4199
 QY 4820 CTCTGAGGGCAGGCTTCTGTGAAAGTTTGCCTGGGCTCAGTGTGAGATTTTGCCATAA 4879
 Db 4200 ----- 4199
 QY 4880 AAGGGTCTCTGCCCTGTGTACAGACAGATCGGAATCTAGAGTGATCTACAGAGTCCC 4939
 Db 4200 ----- 4199
 QY 4940 GCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTTAGATCTCTACGCGGACGCA 4999
 Db 4200 ----- 4199
 QY 5000 TCGTGGCGGGTACCAGAGCTCGAATTCGTAATCATGGTCTAGCTGTTCTGTGTGAAA 5059
 Db 4200 ----- 4199
 QY 5060 TTGTTATCGCTCACAATTCACACACATACAGCGGAGCAATAAGTGTAAAGCCTG 5119
 Db 4200 ----- 4218
 QY 5120 GGGTCCCTAATGAGTGAGTAACTCACATTAATTCGTTGCTGCTCACTGCCCGCTTTCCA 5179
 Db 4219 ----- 4218
 QY 5180 GTCCGGAAACCTGTCTGCGCAGCTCATTAATGATCGGCCAAACGCGGGGAGAGCGG 5239
 Db 4219 ----- 4265
 QY 5240 TTTGGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTG 5299
 Db 4266 TTTGGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTG 4325
 QY 5300 GCTGGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAG 5359
 Db 4326 GCTGGCGGAGCGGTATCAGCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAG 4385
 QY 5360 GGATTAACGAGAAAGACATGTAGAGCAAAAGCCAGCAAAAGGCCAGGAACCGTAAAAA 5419
 Db 4386 GGATTAACGAGAAAGACATGTAGAGCAAAAGCCAGCAAAAGGCCAGGAACCGTAAAAA 4445
 QY 5420 GCGCGGTGCTGGGCTTTTCCATAGCTCCGCCCTGACGAGCATCAAAAATCG 5479
 Db 4446 GCGCGGTGCTGGGCTTTTCCATAGCTCCGCCCTGACGAGCATCAAAAATCG 4505
 QY 5480 AGCTCAAGTCAAGGTGGCGAAACCCGACAGGACTATAAAGATACACAGCGCTTTCCGCC 5539
 Db 4506 AGCTCAAGTCAAGGTGGCGAAACCCGACAGGACTATAAAGATACACAGCGCTTTCCGCC 4565
 QY 5540 TGGAGTCCCTCTGCGCTCTCTGTTCCCGACCTGCGCTTACCGGATACCTGTCGCC 5599
 Db 4566 TGGAGTCCCTCTGCGCTCTCTGTTCCCGACCTGCGCTTACCGGATACCTGTCGCC 4625
 QY 5600 CTTTCTCCCTTCCGGAAGCGTGTCTCATAGCTCACGCTGAGGTATCTCAGTTC 5659
 Db 4626 CTTTCTCCCTTCCGGAAGCGTGTCTCATAGCTCACGCTGAGGTATCTCAGTTC 4685
 QY 5660 GGTGTAGTGTGTTCTGCTCAAGCTGGGTGTGTGCAAGCCCGCTTTCAGCGCCGACCG 5719

```
Db 4686 GGTGTAGTTCGTTCTCAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACGG 4745
QY 5720 CTGCGCTTATCCGGTAACATACTGCTTGGAGTCCAAACCCGGTAAAGACGACGATTAATCGCC 5779
Db 4746 CTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAAGACGACGATTAATCGCC 4805
QY 5780 ACTGCAGACGACCACTGGTTACAGAGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGA 5839
Db 4806 ACTGCAGACGACCACTGGTTAAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGA 4865
QY 5840 GTTCTTTGAAGTGGTGGCTTAACTAGCGGTACACTAGAGGACAGATTATTTGGTATCTGCGC 5899
Db 4866 GTTCTTTGAAGTGGTGGCTTAACTAGCGGTACACTAGAGGACAGATTATTTGGTATCTGCGC 4925
QY 5900 TCTGCTGAAGCCAGTTACTCTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAAC 5959
Db 4926 TCTGCTGAAGCCAGTTACTCTCGGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAAC 4985
QY 5960 CACCGCTGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGG 6019
Db 4986 CACCGCTGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGG 5045
QY 6020 ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTC 6079
Db 5046 ATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACCTC 5105
QY 6080 AGCTTAAGGATTTTGGTCAAGATTATCAAAAAAGGATCTTCACTAGATCTCTTTTAAA 6139
Db 5106 AGCTTAAGGATTTTGGTCAAGATTATCAAAAAAGGATCTTCACTAGATCTCTTTT--- 5162
QY 6140 TTAATAAATGAAGTTTAAATCAATCAATGAATATATATGAGTAAACCTTGGTCTGACAGTTA 6199
Db 5163 ---GCGGCCGCGCCCAATCAATCAATGAATATATATGAGTAAACCTTGGTCTGACAGTTA 5218
QY 6200 CCAATGCTTAATCAGTGAAGGACCTTATCTCAGCGATCTGCTTATTTGGTTTCATCCATAGT 6259
Db 5219 CCAATGCTTAATCAGTGAAGGACCTTATCTCAGCGATCTGCTTATTTGGTTTCATCCATAGT 5278
QY 6260 TGCTGACTCCCGCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCCAG 6319
Db 5279 TGCTGACTCCCGCTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTGGCCCCCAG 5338
QY 6320 TGCTGCAATGATACCGCAGAGCCACGCTACCGGCTCCAGATTATCAGCAATAAACCA 6379
Db 5339 TGCTGCAATGATACCGCAGAGCCACGCTACCGGCTCCAGATTATCAGCAATAAACCA 5398
QY 6380 GCCAGCCGAAAGGCGGAGCGCAGAGTGGTCTCGCAACTTATCCGCTCCATCCAGTC 6439
Db 5399 GCCAGCCGAAAGGCGGAGCGCAGAGTGGTCTCGCAACTTATCCGCTCCATCCAGTC 5458
QY 6440 TATTAATTTGTTGCCGGGAAGCTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTGGCGCAAGCT 6499
Db 5459 TATTAATTTGTTGCCGGGAAGCTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTGGCGCAAGCT 5518
QY 6500 TGTGTCATTTGCTACAGGC-TGTGGGTGTCAGCTCGTGGTTGGTATGGCTTCAATCAG 6558
Db 5519 TGTGTCATTTGCTACAGGCATCGTGGTGTACGCTCGTGGTTGGTATGGCTTCAATCAG 5578
QY 6559 CTCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGT 6618
Db 5579 CTCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGT 5638
QY 6619 TAGTCTCTTCGCTCTCCGATCGTTGTGTCAGAAAGTAAGTTGGCGCAGTGTATCACTCAT 6678
Db 5639 TAGTCTCTTCGCTCTCCGATCGTTGTGTCAGAAAGTAAGTTGGCGCAGTGTATCACTCAT 5698
QY 6679 GGTATGCGACGACTGCATAATTTCTTATCTGTCTATGCCATCCGTAAGATGCTTTTCTGT 6738
Db 5699 GGTATGCGACGACTGCATAATTTCTTATCTGTCTATGCCATCCGTAAGATGCTTTTCTGT 5758
QY 6739 GACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTC 6798
Db 5759 GACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTGTATGCGGCGACCGAGTTGCTC 5818
```

```
QY 6799 TTGCCCCGGCTCAATAACGGGATAATAACCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCAT 6858
Db 5819 TTGCCCCGGCTCAACACGGGATAATAACCGCGCCACATAGCAGAACTTTAAAAAGTGTCTCAT 5878
QY 6859 CATTGGAAAAAGTCTTTCGGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG 6918
Db 5879 CATTGGAAAAAGTCTTTCGGGGCGGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG 5938
QY 6919 TTCCGATGTAACCCCACTCTCGTGCAACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGT 6978
Db 5939 TTCCGATGTAACCCCACTCTCGTGCAACCCCACTGATCTTTCAGCATCTTTTACTTTTACCAGCGT 5998
QY 6979 TTCTGGGTGAGCAAAAAACAGGAAGCAAAAAATGCGCAAAAAAGGAATAAGGGCCGACAG 7038
Db 5999 TTCTGGGTGAGCAAAAAACAGGAAGCAAAAAATGCGCAAAAAAGGAATAAGGGCCGACAG 6058
QY 7039 GAAATGTTGAATACTCATACTCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTA 7098
Db 6059 GAAATGTTGAATACTCATACTCTTCTCTTTTCAATATTTTGAAGCATTTTATCAGGGTTA 6118
QY 7099 TTGTCTCATGAGCGGATACATATTTGAATGATTTTAGAAAAATAAAATAGGGGTTC 7158
Db 6119 TTGTCTCATGAGCGGATACATATTTGAATGATTTTAGAAAAATAAAATAGGGGTTC 6178
QY 7159 GCGCACATTTCCCGAAAAAGTGCCACCTGACGTCTAAGAAACCATTTATTCATGACATT 7218
Db 6179 GCGCACATTTCCCGAAAAAGTGCCACCTGACGTCTAAGAAACCATTTATTCATGACATT 6238
QY 7219 AACCTATAAATAATAGGCGTATCAGGAGCCCTTTTCGTC 7256
Db 6239 AACCTATAAATAATAGGCGTATCAGGAGCCCTTTTCGTC 6276

RESULT 9
AB041928 LOCUS AB041928 6248 bp DNA circular SYN 21-MAY-2003
DEFINITION Retroviral vector pCXneo DNA, complete sequence.
ACCESSION AB041928
VERSION AB041928.1 GI:7649702
KEYWORDS
SOURCE Retroviral vector pCXneo
ORGANISM Retroviral vector pCXneo
other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS Akagi, T., Shishido, T., Murata, K. and Hanafusa, H.
TITLE v-Crk activates the phosphoinositide 3-kinase/AKT pathway in transformation
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7290-7295 (2000)
MEDLINE 20319019
PUBMED 10852971
REFERENCE 2 (bases 1 to 6248)
AUTHORS Akagi, T.
TITLE Direct Submission
JOURNAL Submitted (17-APR-2000) Teiyoshi Akagi, Osaka Bioscience Institute, The First Department; 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan (E-mail: takagi@obi.osaka-u.ac.jp, Tel: 81-6-6872-4850, Fax: 81-6-6871-6686)
FEATURES
source Location/Qualifiers
1..6248 /organism="Retroviral vector pCXneo"
/mol_type="genomic DNA"
/db_xref="taxon:197519"
enhancer 1..466 /note="CMV enhancer"
467..762 /note="R-U5 region of MolU 5'LTR"
763..1673 /note="psi+; MolU extended packaging signal"
1674..2049 /note="MolU pol region containing splice acceptor site"
2052..2136 /note="multiple cloning site"
```


QY 2461 GCACAGCTACAGGTGTCCACTCCGACATCCAGTCCAGCCAGAGCCCAAGCAGCCTGAGC 2520
Db 2084 GGTACCCGGGGCGCGCGATCTTAAATTAAGCGCGCGGTAACTATTGTTAACTAAC 2143
QY 2521 -----GCCAGCGTGGGTGACAGAGTACCATCACCTGTAAAGCCAGTCAAGGATGTGGG 2573
Db 2144 TAACTAGCTAGCAACGGTTTCCCTCTAGCGGGATCAATTCGGCCCCCCCCCTAACGT 2203
QY 2574 TACTTCTGTAGTGTGTACAGCAGAGAACCCAGGTAAAGGCTCAAAGCTGTGATCTACTG 2633
Db 2204 TACTGGCCGAAGCGCTTGGAAATAAGCGCGGTGCGTGTCTCTATATGTTATTTTCCAC 2263
QY 2634 GACATCCACCC-----GGCACACTGGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTA 2689
Db 2264 CATATGCGCTCTTTTGGCAATGTGAGGGCCGGAACCTGGCCCTGTCTTCTTTGACGAG 2323
QY 2690 CCGACTTCACCTCAGCAGCCTCCAGCCAGAGGACATCGCCACCTACTACTGCCC 2749
Db 2324 CATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCGTGAA 2383
QY 2750 AGCAATATAGCTCTATCGGTGTTCGCCCAAGGGAACCAAGGTGGAATCAAAACGAGGTG 2809
Db 2384 GGAAGCAGTCTCTCTGGAAGCTTCTTGAAGACAAACACAGCTCTGTAGCGACCTTTTGCAG 2443
QY 2810 GCTCAGGATCGGTGGATCCGGCTCTGTGCTCAGGATCGAGGTCCAACCTGTGGAGA 2869
Db 2444 GCAGCGGAACCCCCACCTCGCGCAGCAGGTCTCTCGCGCCAAAGCAGCAGTGTATAAGA 2503
QY 2870 -----GCGGTGGAGTGTGTGCAACCTGCGCGGTCTCCTGCGCTCTCTCGCATC 2924
Db 2504 TACACTGCAAGCGGCAACCCAGTGCACGTTGTGTGATGTGATGTGTGGAAG 2563
QY 2925 TGGCTTCGATTTCACCATATTTGGATGAGTTGGGTGAGACAGGCACCTCGTGAAGGTCT 2984
Db 2564 AGTCAATGGCTCTCTCAAGCGTATTCAACAGGGCTGAAGGATGCCAGAGGTACC 2623
QY 2985 TGAGTGGATTTGGAAATATTCACAGATAGCAGTACAGTAACTATGCGCGCTCTCTAAA 3044
Db 2624 CCATTGTATGGATCTGATCTGGGGCCTCGGTGCAC-ATGCTTTACGTGTGTTTAGTGA 2682
QY 3045 GGATAGATTTTACATATTCGCGAGACAGCCCAAGACACATTTGTTCTGCAAA----- 3097
Db 2683 GGTAAATAAAGCTTAGGCCCCCGAACCAGGGGACGTGTTTCTCTTTGAAATAACAG 2742
QY 3098 TGGACAGCTCAGACCCGAGACACCGGGTCTATTTTGTGCAAGCCTTTACTTTCGGCT 3157
Db 2743 ATATACCATGATTGAACAAGATGATTGACGAGGTTCCTCCGGCGCTTGGGTGAGA 2802
QY 3158 TCCCTGTGTTGCTTATTTGGGGCCAGAGGACCCCGGTCAACGCTCCAGTGTCTAAGCCCA 3217
Db 2803 GGCTATTCCGCTATGACTGGGCAACAACAGACAATCGGTGCTCTGATGCCCGGTGTTCC 2862
QY 3218 CCACGAGCCAGCCCGGACCAACAACCGCGCCACCATCGCTCGCAGCCCTGT 3277
Db 2863 GGCTGTACGCGAGGGGCGCCGGTCTTTTGTGCAAGCCGACTGTGCGGGTCCCTGA 2922
QY 3278 CCCTGCGCCACAGAGCGG-----CTCGCCAGCGGGGGGGGGCGAGTGACA 3325
Db 2923 ATGAACTGCAGACGAGCGCGGTATCGTGGCTGGCCACAGCGGGCTTCTTTGG 2982
QY 3326 CGAGGGGCTGGATTCGCCCTGGATCCAAACTCTGCTACTCTGATGGAATCTCT 3385
Db 2983 CAGCTGTGCTCGAGTGTGCTACTGAAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC 3042
QY 3386 TCATCTATGGTGTCTATCTACTGCTTCTGAGAGTGAAGTTACAGAGGAGCGCAG 3445
Db 3043 CGGGGCGGATCTCCTGTCTATCTACCTTGCTCTGCGCGAAGATATCCATCATGCTG 3102
QY 3446 AGCC-----CCCCCGGTATACAGCAGGGCCAGAACAGCTCT 3481
Db 3103 ATGCAATGCGCGGCTGATAGCTTGTATCGGCTACTGCGCCATTGACCAACAGCGA 3162
QY 3482 ATAAAGAGCTCAATCTAGGACGAAGAGAGGAGTACGATGTTTGGACAAGAGAGCTGGCC 3541

Db 3163 AACATCGATCGAGCGAGCAGTACTCGATGGAAGCCGGTCTTGTGATCAGGATGATC 3222
QY 3542 GGGACCTTGAGATGGGGGAAAGCCGAGAAAGAAAGAAACCTCTCAGAAAGCCCTGTACAAATG 3601
Db 3223 TGGAGGAAGAGCATCAGG--GCTCGCGCCAGCCGAACTGTTCCGCCAGCTCAAGCGCG 3280
QY 3602 AACTGCAAGAAAGATAAGATGGCGGAGCCTTACAGTGAGATTGGGATGAAAGCGGAGCGCC 3661
Db 3281 CATGCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCCCTGCTTGCAGAAATATCAT 3340
QY 3662 GGAGGGGCAAGG--GGCACGATGGCCTTTACCAGGTCTCAGTACAGCCACCAAGGACAC 3719
Db 3341 GGTGAAATATGCGCGCTTTTCTGGAATTCATGACTGTGGCGGCTGGGTGCGGACCG 3400
QY 3720 CTACGAGCGCCTTCACATGACGAGCGCTCGCCCTCGCTAA-----CTGACGCGCGCGG 3775
Db 3401 CTATCAGACATAGCGTGGCTACCGCTGATATTTGCTGAGAGCTTGGCGGGAATGGG 3460
QY 3776 ATCCGGAATTAGTCAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGAC 3835
Db 3461 TGACCGCTTCTCTGTGCTTTACGGTATCGCGCTCCCGATTCGCGAGCGCATCGCCTTCTA 3520
QY 3836 TCAACATATACACAGCTGAAGCCTATAGAGTACGACCATAGATAAATAAAGATTTT 3895
Db 3521 TCGCCTTCTTGAAGAGTTCTTGAAGCGGACTCTGGGGTTTCGATAAATAAAGATTTT 3580
QY 3896 ATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGC 3955
Db 3581 ATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTTTGGCAAGCTAGC 3640
QY 3956 TTAAGTAAACCCATTTTTCAGGCGATCG--AAAATACATACTGAGATAGAGAGTTTCA 4014
Db 3641 TTAAGTAAACCCATTTTGGAGGCGATGGAATAATACATACTGAGATAGAGAGTTTCA 3700
QY 4015 ATCAGGTTAGGAACAGAG--ACAGCAGATATATGGGCCAAACAGGATATCTGTGTTAG 4073
Db 3701 ATCAAGTTCAGGAACAGATGGAACAGCTGAAATATGGGCCAAACAGGATATCTGTGTTAG 3760
QY 4074 CAGTTCCTGCCCC--GCTCAGGGCCAAAGAAAGTTGGAACAGGAGAAATATGGGCCAAACAG 4132
Db 3761 CAGTTCCTGCCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAACAG 3820
QY 4133 GATATCTGTGTAAGCAGTTCTCTGCCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGAT 4192
Db 3821 GATATCTGTGTAAGCAGTTCTCTGCCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGAT 3880
QY 4193 GCGTCCCGCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCCCCAAGGA 4252
Db 3881 GCGTCCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTCCCCCAAGGA 3940
QY 4253 CCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCAGTTCGCTTCTGCTTCTGTTTC 4312
Db 3941 CCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCAGTTCGCTTCTGCTTCTGTTTC 4000
QY 4313 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCTCTACTCGCGCGCGCAG 4372
Db 4001 GCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCTCTACTCGCGCGCGCAG 4060
QY 4373 TCCTCCGATAGCTGCTGCTGCGGTACCGGTGTTCTCAATAAACCTCTTTGAGTTGC 4432
Db 4061 TCCTCCGATGACTGAGTCCCGGGTACCGGTGATATCAATAAACCTCTTTGAGTTGC 4120
QY 4433 ATCCGACTCTGCTGCTGCTGCTTCCCTTGGGAGGGTCT-CTCTGATGATGACTACCGGT 4491
Db 4121 ATCCGACTCTGCTGCTGCTGCTTCCCTTGGGAGGGTCTCTCTCTGATGATGACTACCGGT 4180
QY 4492 CAGCGGGTCTTTCAGTTCCTCCACCTACACAGGTCTCACTAACTTCTCTGATGTGCGG 4551
Db 4181 CAGCGGG----- 4188
QY 4552 CAGGACTCCTGTCAGCCCGGTTTTTGTATATAAATAAGAAACAGTGTTCCTTCA 4611

Db 4189 ----- 4188
 QY 4612 AGCCAGACTACATCTCTCGGCTTTATATAAGATGTTGAAGGCTCTGTGGACTA 4671
 Db 4189 ----- 4188
 QY 4672 TCTGCCACAGACTTTTAAAGATTTTATATGCTCTCTGGATGAGGATTTAGTCAATCTAT 4731
 Db 4189 ----- 4188
 QY 4732 CCTCGTCTAATTTGTCTGGCTTCTCCGATTTTAAATTTCTAGTTGGACTCCCTTCCTGA 4791
 Db 4189 ----- 4188
 QY 4792 GAGCAGCGGATTCAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTGCC 4851
 Db 4189 ----- 4188
 QY 4852 TGGGCTCAGTGTGAGATTTTGGCATAAAAAAGGGTCTGCCCCCTGTGTACAGACAGATCG 4911
 Db 4189 ----- 4188
 QY 4912 GAATCTAGAGTGCACTACTCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCATCTT 4971
 Db 4189 ----- 4188
 QY 4972 TGCCTAGAGATCCTCTACGCCGAGCGCATCGTGGCGGGTACCGAGCTCGAATTCGTAAT 5031
 Db 4189 ----- 4188
 QY 5032 CATGCTCATAGCTGTTTCTCTGTGAAATTTGTTATCCGCTCACAATTTCCACACACATAC 5091
 Db 4189 ----- 4188
 QY 5092 GAGCGGGAAGCATAAAGTGTAAAGCTGGGGTGCTTAATGAGTGAGCTAACTCACATTA 5151
 Db 4189 ----- 4188
 QY 5152 TTGGCTTGCGCTCACTGCCCGCTTTCCAGTCGGGAACCTGTGCGCAGCTGCATTAAT 5211
 Db 4189 ----- 4188
 QY 5212 GAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTCCGCTTCTCTGCG 5271
 Db 4209 GAATCGGCCAACCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTCCGCTTCTCTGCG 4268
 QY 5272 TCACTGACTCGCTGCGCTCGGTCGTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAGG 5331
 Db 4269 TCACTGACTCGCTGCGCTCGGTCGTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAGG 4328
 QY 5332 CGGTAATACGGTTATCCACAGATCAGGGGATAACGACGGAAGACATGTGAGCAAAAG 5391
 Db 4329 CGGTAATACGGTTATCCACAGATCAGGGGATAACGACGGAAGACATGTGAGCAAAAG 4388
 QY 5392 GCCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTGTGCGGTTTCCATAGGCTCC 5451
 Db 4389 GCCAGCAAAAGGCCAGGAACCGTAAAGGCGCGTGTGCGGTTTCCATAGGCTCC 4448
 QY 5452 GCCCCCTCAGCAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGGAACCCGACAG 5511
 Db 4449 GCCCCCTCAGCAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGGGAACCCGACAG 4508
 QY 5512 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTCCGA 5571
 Db 4509 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCGA 4568
 QY 5572 CCTGCGCTTACCGGATACCTGCTCGGCTTTCTCCCTTTCCGGAAGCGTGGGCTTCTC 5631
 Db 4569 CCTGCGCTTACCGGATACCTGCTCGGCTTTCTCCCTTTCCGGAAGCGTGGGCTTCTC 4628
 QY 5632 ATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTAGGTGTAGGTGTAGGTGTAG 5691
 Db 4629 ATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTAGGTGTAGGTGTAGGTGTAG 4688

QY 5692 TGCAGAACCCCGCTTTCAGCCCGCTGCGCTTATCCGGTAACTATCCTCTTGAGT 5751
 Db 4689 TGCAGAACCCCGCTTTCAGCCCGCTGCGCTTATCCGGTAACTATCCTCTTGAGT 4748
 QY 5752 CCAACCCGGTAAAGACACGACTTATCGGCACCTGGGACGACCTGGTAAACAGGATTAGA 5811
 Db 4749 CCAACCCGGTAAAGACACGACTTATCGGCACCTGGGACGACCTGGTAAACAGGATTAGA 4808
 QY 5812 GAGCAGGATGTAGGCGGTCTACAGAGTCTTGAAGTGTGGCTAACTACGGCTACA 5871
 Db 4809 GAGCAGGATGTAGGCGGTCTACAGAGTCTTGAAGTGTGGCTAACTACGGCTACA 4868
 QY 5872 CTAGAAGACAGTATTTGGTATCTGCGCTCTGCTCAAGCCAGTATACCTTCGGAAGGAG 5931
 Db 4869 CTAGAAGACAGTATTTGGTATCTGCGCTCTGCTCAAGCCAGTATACCTTCGGAAGGAG 4928
 QY 5932 TTGGTGTCTCTGATTCGGCAAAACAAACCGCTGTAGGCGGTGTTTTTGTGTTGA 5991
 Db 4929 TTGGTGTCTCTGATTCGGCAAAACAAACCGCTGTAGGCGGTGTTTTTGTGTTGA 4988
 QY 5992 AGCAGCAGTAAACCGCAGAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGG 6051
 Db 4989 AGCAGCAGTAAACCGCAGAAAGGATCTCAAGAGATCCTTTGATCTTTCTACGG 5048
 QY 6052 GGTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTAGATTAATCAA 6111
 Db 5049 GGTCTGACGCTCAGTGGAAACGAAACTCAAGTAAAGGATTTTGGTCTAGATTAATCAA 5108
 QY 6112 AAAGATCTTCACTAGATCTCTTTTAAATTAAGTAAAGTAAATCAATCTAAAGTA 6171
 Db 5109 AAAGATCTTCACTAGATCTCTTTTAAATTAAGTAAAGTAAATCAATCTAAAGTA 5161
 QY 6172 TATATGATTAACCTGGTCTGACAGTAACTCAATGCTTAATCAGTGAGCAGCTATCTCAG 6231
 Db 5162 TATATGATTAACCTGGTCTGACAGTAACTCAATGCTTAATCAGTGAGCAGCTATCTCAG 5221
 QY 6232 CGATCTGTCTATTTTCGTTTCATCCATAGTGTGCTGACTCCCGCTCGTGTAGATACTAGA 6291
 Db 5222 CGATCTGTCTATTTTCGTTTCATCCATAGTGTGCTGACTCCCGCTCGTGTAGATACTAGA 5281
 QY 6292 TACGGGAGGCTTACCATCTGCGCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCAC 6351
 Db 5282 TACGGGAGGCTTACCATCTGCGCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCAC 5341
 QY 6352 CGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGGCGGAGCGGAGAGTGTGTC 6411
 Db 5342 CGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGGCGGAGCGGAGAGTGTGTC 5401
 QY 6412 CTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTA 6471
 Db 5402 CTGCAACTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTA 5461
 QY 6472 GTTCCGCAAGTAAATAGTTTGGCAGACGTTGTGCGCAATGCTGCTACAGGC-TGCGTGTGTCAC 6530
 Db 5462 GTTCCGCAAGTAAATAGTTTGGCAGACGTTGTGCGCAATGCTGCTACAGGCATCGTGTGTCAC 5521
 QY 6531 GCTCGCTGTTGGTATGCTTTCATTCAGTCCGTTCCCAAGATCAAGCGGAGTTACAT 6590
 Db 5522 GCTCGCTGTTGGTATGCTTTCATTCAGTCCGTTCCCAAGATCAAGCGGAGTTACAT 5581
 QY 6591 GATCCCCCATGTTGTGCAAAAGCGGTTAGTCTCTCGGTCTCTCGATCGTGTGTAGAA 6650
 Db 5582 GATCCCCCATGTTGTGCAAAAGCGGTTAGTCTCTCGGTCTCTCGATCGTGTGTAGAA 5641
 QY 6651 GTAAAGTTGGCGCAGTGTATCACTCATGTTTATGCGAGCATGATTAATCTCTTACTG 6710
 Db 5642 GTAAAGTTGGCGCAGTGTATCACTCATGTTTATGCGAGCATGATTAATCTCTTACTG 5701
 QY 6711 TCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTCTCAACCAAGTCACTCTGAG 6770
 Db 5702 TCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTCTCAACCAAGTCACTCTGAG 5761

1365	DB	GAATGGCCAACTTTAACTGTCGATGGCCGGGAGACGGACACCTTTTAAACCGAGACCTCATC	1424
1806	QY	ACCAGGTTAAGATCAAGGTCCTTTTTCACCTGGGCCGCGATGGACACCCAGACCAAGGTC	1865
1425	DB	ACCAGGTTAAGATCAAGGTCCTTTTTCACCTGGGCCGCGATGGACACCCAGACCAAGGTC	1484
1866	QY	TACATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCGCCCTCCCTGGGTCAAGCCCTTGTGA	1925
1485	DB	TACATCGTGACCTGGGAAGCCTTGGCTTTTGTGACCCGCCCTCCCTGGGTCAAGCCCTTGTGA	1544
1926	QY	CACCTAAGCCTCGGCCTCTCTCTTCATCGCCCGCCGCTCTCTCCCTTTGAACCTCTCT	1985
1545	DB	CACCTAAGCCTCGGCCTCTCTCTTCATCGCCCGCCGCTCTCTCCCTTTGAACCTCTCT	1604
1986	QY	CGTTGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTCTCTAGGCGGCC	2045
1605	DB	CGTTGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTCTCTAGGCGGCC	1664
2046	QY	ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTCTGTAACTTCCCTGACCT	2105
1665	DB	ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTCTGTAACTTCCCTGACCT	1724
2106	QY	GACATGACAGAGTTACTACAGCCCTCTCTCCAGCTCACATTACAGGCT-TCTACTTGA	2164
1725	DB	GACATGACAGAGTTACTACAGCCCTCTCTCCAGCTCACATTACAGGCTCTCTACTTGA	1784
2165	QY	GTCCAGCACAAAGTCTGGAGACCTCTGGCGGACGCTTACCAAGAACTGGACCCGACCG	2224
1785	DB	GTCCAGCACAAAGTCTGGAGACCTCTGGCGGCA-CGTACCAAGAACATGGACCGACG	1843
2225	QY	GTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGTCCGCGACACCAAGACTAAG	2284
1844	DB	GTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGTCCGCGACACCAAGACTAAG	1903
2285	QY	AACCTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGTGACCAACCCCAACCGCCTC	2344
1904	DB	AACCTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGTGACCAACCCCAACCGCCTC	1963
2345	QY	AAAGTAGACGGCATCGCAGCTTGATACACGCGCCACCTGAAGGTGCGGACCCGCGGG	2404
1964	DB	AAAGTAGACGGCATCGCAGCTTGATACACGCGCCACCTGAAGGTGCGGACCCGCGGG	2023
2405	QY	GGTGGACCATCCTTAGACTGCCATGGGATGGAGCTGTATCATCTCTTCTTGTTAGCAA	2464
2024	DB	GGTGGACCATCCTTAGACTGCCGATCGAATGGTACCGAGCTCGATC-CACTAGTAA	2082
2465	QY	CAGTTACAGTGTCCATCCGACATCAGCTGACCCAGAGCCCAAGCAGCCTGAGCGCCA	2524
2083	DB	CGGCGGCAGTGTGCTGGAGCCCATCTCTATTATAAAATGTCTCAGAGCAACCGGGAGC	2142
2525	QY	CGGTGGGTGACAGATGACG-ATCACCTGTAAAGCCAGTCAGGATGTGGGTACTTCTGTGA	2583
2143	DB	TGGTGGTGTGACTTCTCTCTTCAAGCTCTCCAGAAAGGATACAGCTGGAGTCAGTTTGA	2202
2584	QY	GCTTTGTATCAGCAGAAAGCCAGGTAAGGCTTCAAAGCTGTGTATCTACTGGACATCCACC	2643
2203	DB	GCGATGTGGAAGAAACAGGACTGAAGCCCAAGAGAAATCTGAACCAAGAAAGGAGACCC	2262
2644	QY	CGGACACTGGTGTGCCAAGCAGATTCAAGCGGTGACGGGTAGCGGTACCGACTTCACTTC	2703
2263	DB	CCAGTGCATCAATGGGCAACCCATCTCTGGCACCTTGGCGGATAGCCCCCGGTGAATGGAG	2322
2704	QY	ACCATCAGCAGCCTCCAGCCAGGACATCGCCACCTACTACTCTCCAGCAATATAGCCTC	2763
2323	DB	CCACTGCCCAACAGCAGTGTGGATGCGCGGGAGGTACTCCCAATGGCAGCAGTGAAGC	2382
2764	QY	TATCGGTGTGCTCGGCCAAGGGAACCAAGT-----GGAAATCAAACGAGGTGG	2810
2383	DB	AAGCGCTGAGAGGCTGGCGATGAGTTTGNACTGCGGTACCGGAGAGCAATCAGTGATC	2442
2811	QY	CTCAGGATCGGGTGGATCCGGCTCTGTGGCTCAGGATCGGAG-GTCCAACTGGTGGAGA	2869
2443	DB	TAACTCCCAAGCTTCATATAACCCAGGACAGTATATCAGAGCTTTTGAACAGGTAGTGA	2502

Qy	2870	CGCGTGAAGTGTTGTGCAACCTGGCCGGTCCCTGCGCCTGTCTCTGCTCGCATCTGGCT	2922
Db	2503	ATGAACCTCTTTTCGGGATGGGGTAAACTGGGGTCGCAATTGFGGCTCTCTCTCTTTGGGG	2562
Qy	2930	TCGATTTCCACCATATTTGGATGAGTTGGGTGAGACAGGCACCTGGAAAAAGGTCTTTGAGT	2989
Db	2563	GGGCACTGTGCTGGAAAGCGTATAGACNAGAGAGATGACGGTATTTGGTCAGTCCGATTGCAA	2622
Qy	2990	GGATTGGGAAAAATTCAATCCAGATAGCAGTACGATTAACTATGCGCCGTCTCTAAAGGATA	3049
Db	2623	G--TTGGATGGCCACCTTACCTGAATGACCACCTAGAGCCTTGGATCCAGGAGAACGGCGG	2680
Qy	3050	GATTTACAATATCGCGAGACAAAGCGCAAGAACATTTGTTCTGCAAAATGGACAGCCCTGA	3109
Db	2681	CTGGGACACCTTTTGTGGATCTTACGGGAAACAATACAGCACCCGAGAGCCGGAAGGCCCA	2740
Qy	3110	GACCCGAAGACACCGGGGCTCTATTTTTGTGCAAGCCCTTTACTTCGGCTTCCCTCGTTTG	3169
Db	2741	GGAGCGTTTCAACCGCTGGTTCTTGACGGGCATGACTGTGGCTGGTGTAGTTCTGTCTGGG	2800
Qy	3170	CTTATTTGGGCCCCAAGGAGCCCCGGTACCGTCTCCAGTGTCTAGCCCAACACAGCGCCAG	3229
Db	2801	CTCACTCTTCAGTCCGGAAGTGACCAAGTTCTGCAGATATCCATCACACTGGCGGCGGAA	2860
Qy	3230	CGCGCGACCAACAAACCGCGGCCCA-----CCATCGCGTCGCGAGC	3271
Db	2861	TTCCGGATCCACTAGTAAACCGCGCCAGTGTGTGGAATTCGCCCCCTCTCCCTCCCCC	2920
Qy	3272	CCCTGTCCCTGCGCCCAAGAGCGCGCTGGCCAGCGGGGGGCGCAGTGCACACGAGGG	3331
Db	2921	CCCTAAAGTTACTTGGCCGAAGCGCTTGGAAATAAGGCCGGTGTGTGTTGTCTATATGTG	2980
Qy	3332	GGCT-----GGACTTCGCGCCTCGATC	3352
Db	2981	ATTTTCCACATATGCGGTCTTTTGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTT	3040
Qy	3353	CCAAACTCTGCTACTCTCGATGGAATCTCTTCTATGTGTCTCATTTCTCACTGCGCT	3412
Db	3041	CTTGACGAGCATTCCTAGGGGTCTTCCCTCTCGCCAAAGGAATCAAGGTCTGTTGAA	3100
Qy	3413	TGTTCTCGAGAGTGAAGTTCAGAGG-----AGCCAGAGCCCCC	3453
Db	3101	TGTCGTGAAGGAAGCAGTTCTCTGGAAGCTCTTGAAGACAAACACGCTCTGTAGCGAC	3160
Qy	3454	CGCTACCAGCAGGGCCAGAACCAACAGCTCTATAACGAGCTCAATCTAGGACGAAGAGAGAG	3513
Db	3161	CTTTTGACGGCAGCGGAAACCCCCACCTGGCGACAGTGCCTCTGCGGCCAAAAGCCACG	3220
Qy	3514	TACGATGTTTTTGGACAAGACGCTGGCCGGGACCTTGAGATGGGGGAAAGCCG-----	3567
Db	3221	TGTATAAGATACACTCTCAAAGCGGCACAAACCCAGCTGCCAGTTGTGTAGTTGGATAGT	3280
Qy	3568	-----AGAAGNAGAACCTCAGGAGGCCCTGTAAATGAACTGCAGAAAGATAAGAT	3620
Db	3281	TGTGGAAAGATCAAAATGGGTCTCTCTCAAGCGTAGTCAACAGGGGCTGAAGGATGCCCA	3340
Qy	3621	GGCGGAGGCTTACAG--TGAGATTGGATGAAAGGC----GAGCGCCGAGGGGCAAGGG	3674
Db	3341	GAAGGTACCCCATTTGTATGGGAATCTGATCTGGGGCCTCGGTGCAATGCTTTTACATGTG	3400
Qy	3675	GCACGATGGCCTTATCCAGGGTCTCAGTACAGC--CACCAGAGCACCTTACGACGCCCTTC	3733
Db	3401	TTTATGTCGAGGTTAAAAAAGCTCTAGGCCGCCCGGAACACCGGGGACGTGGTTTTTCTCTTG	3460
Qy	3734	ACATGAGGCCCTCGCCCTCTCGTAACTCGACGGCGCGCGATCCGGATTTAGTCCCAATT	3793
Db	3461	AAAAACAGATGATPAAGCTTGCCACAGGAATTCATGTGTAGCAGCAGATCCTCAAGAAC	3520
Qy	3794	TGTTAAAGACAGGATATCAGTGG--TCCAGGCTCTAGTTTTTGACTCAACAATATCACAG	3851
Db	3521	ACGGGCTTGAGGAGATCATGTAGCTTCAAGGTGAACCTGGAGGCGGTGGTGAACAACGAC	3580

Qy	3952	CTGAAGCCTATAGAGTAC - GAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAA	3911
Db	3581	GTGTTCCACATGGAGGCTCGGCAAGGGCAACATCTCTGTTCCGCAACCAAGCTGGTGCAG	3640
Qy	3911	AAAGGGGGGAATGAAGACCCACCTGTAGGTGTTGGCAAGCTAGCTTAAGTAAAGCCCATTT	3970
Db	3641	ATCCGCGTGACCAAGGGGCCCCCTGCCCCTTGACATCTCGAGCCCGCCCTTC	3700
Qy	3971	TTGCAAGGCA ----- TGGAAATAATATAACTGAGAATAGAG	4006
Db	3701	CAGTACGGCAACCGCACCTTCACCAAGTACCCCGAGGACATCAGCGACTTCTTCATCCAG	3760
Qy	4007	AAGTTCAAGTACAGGTTAGGAACAGAGAGACAGAGAAATAGGGCCAAAACAGGATATCTG	4066
Db	3761	AGCTTCCCGCGGCTTCGTGTACAGCGCACCCCTGGCGCTACGAGGACGGCGCGCTGGTG	3820
Qy	4067	TGG ----- TAAGCAGTTCCTGCCCCGCTCAGGGCCAAACAGTGTGAACA	4112
Db	3821	GAGATCCGACGACATCACTGATCGAGAGATGTTGTTATCCCGTGGAGTACAAG	3880
Qy	4113	GGAGAATATGGGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGCGCTCAGGGCCA	4172
Db	3881	GGCGCAACTTCCCAACGACGGCCCCGTGATGAAGAAGCAACATCAGCGCTCGACGCC	3940
Qy	4173	AGAACAGATGTTCCCAAGATGCGGTCCGCGCTCAGCAGTTCCTAGAGAAC -	4223
Db	3941	AGCTTCGAGGTGGTGTATATGAACGACGGCGTGCTGCTGGGCCAGGTGATCTGGTGTAC	4000
Qy	4224	----- CATCAGATGTTTCAGGGTCCCCCAAGGACCTGAAATGAC	4263
Db	4001	CGCTGAAACAGCGCAAGTCTACAGCTGGCACATGCGCACCTTGATGAAGAGCAAGGGC	4060
Qy	4264	CCTGTGCTTATTTTGAACATA - ACCAATCAGTTCGCTTCTCGTTCCTGTTTCGGCGCGTTC	4321
Db	4061	GTGTTGAAGACTTCCCGAGTACCACTCATCCAGCACCGCTCGGAGAGAACCTAGCTG	4120
Qy	4322	TGTCCTCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGC -	4369
Db	4121	GAGGACGGCGCTTCGTGGAGCAGCAGAGACCGCCATCGCCAGCTGACCAAGCTGGGC	4180
Qy	4370	CAGTCCTCCGATAGACTGCGTGGCGCGGTACCGTGTCTCAATAAACCTCTTGCAGT	4429
Db	4181	AAGCCCCCTGGCGAGCTCGACAGTGGGTGTAACTCGAGCGGCGCGATCCGAGTTCCTC	4240
Qy	4430	TGCATCCGACTCGTGTGCTCGCTTCTTGGGAGGCTCTCT -	4475
Db	4241	TGACGGGACTCTCGGGGTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGG	4300
Qy	4476	GTGATGACTACCCGTGAGCGGGTCTTTGAGTTTCTCCACCTPACACAGGTCTCACTAA	4535
Db	4301	GGAATGAAGAGACCCCACTGTAGGTTTGGCAAGCTAGCTTTAAGTAAAGCCATTTGCAAG	4360
Qy	4536	CATTCTGTATGTGC - CGCAGGAGTCCGTGAGCGCGGTTTGTGTTTATATAAATGCA	4593
Db	4361	GCAATGGAANAATACATAACTGAGAATAGAGAAGTTCAGATCAAGGTTCAGAAACAGATGGA	4420
Qy	4594	AGAACAGTGTTCCTTCAAGCCGACTACATCTCTGACTCTCGGCTTTATAAAGAATGTT	4653
Db	4421	ACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGCAGTTCCTGCCCGCGCTCAGGG	4480
Qy	4654	GAAGGGCTCTGTGGACTATCTGCCACAGCACTTTTAAAGATTTTATG -	4701
Db	4481	CCAAGAACAGATGGAAACAGCTGAATAATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCC	4540
Qy	4702	CCTCCTCGATGAGGGATTTAGTCAATCTATCTCTGCTCTATTTTGTCTGGCTTCTCCGTA	4761
Db	4541	TGCCCGGCTCAGGGCCAAAGACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTT	4600
Qy	4762	TTAATATTC -----	4770
Db	4601	CTAGAGAACCAATCAGATGTTTTCAGGGTGCCCCAAGGACCTGAAATGACCTGTGCTTAA	4660
Qy	4771	----- TAGTTTGCACTCCCTTCTGTAGAGACACGCGCATTCGACAGTAGTT	4815

Db	4661	TTTGAACTAACCAATCAGTTTCGGTTCCTCGCTTCTGCTTCGGCGCGCTCTGCTGCCCGAGCT	4720
Qy	4816	AATACTCTGAGGCGAGCGCTTCTGTGAAAAAGTTGCTCGGGCTCAGTGTGAGATTTTGGCCA	4875
Db	4721	CAATAAAGAGCCACAAACCCCTCACTCGGGCGCCAGTCCCTCCGATTGACTGAGTCGCC	4780
Qy	4876	TAAAAAGGGGTCTGCCCTGTGTGTAACAGACAGATCGGAATCTAGATGCANATCTCAGAGT	4935
Db	4781	CGGGTACCCGTGTATCCAAATAAACCCCTCTTGCAGTTGTCATCCGACTTGTGGTCTCGCTGT	4840
Qy	4936	CCCGCGGTTCCGGGGCTCTGA-----	4957
Db	4841	TCCTTGGGAGGGTCTCCTCTGAGTGATGACTACCGCTCAGCGGGGTCTTTCATTGGG	4900
Qy	4958	-----TCTCAGGGCATCTTTGCTCTAGAGATCTCTACGCCGCGA--GCATCGTG	5004
Db	4901	GGCTCGTCCGGATCGGAGACCCCTGCCACGAGCACCGACCCACCGGAGGTAA	4960
Qy	5005	GCGGGTACCGAGTCGAATTCGTATCATGTGTATAG--CTGTTTCTGTGTGAATTTGT	5063
Db	4961	GCTGGCTGCCCTCGCGGCTTTCGGTGTATGACGGTGAATAACCTCTGACACATGCAGCTCCCG	5020
Qy	5064	TATCCGCTCACAATTCCACACACATACGAGCCGGAAGCATTAAGTGTAAAGCCTGGGT	5123
Db	5021	GAGACGGTCAACGCTTGTCTGTAAAGCGGATGCCCGGAGCAGACAAGCCGTCAGGCGCG	5080
Qy	5124	GCCTAATGAGT-----GAGCTAACTCA	5145
Db	5081	TCAGCGGGTGTGGCGGGTGTGGGGCGCAGCCATGACCCAGTCACGTAGCATGCGGA	5140
Qy	5146	CATTAAATGCGTTGCGCTCATCTGCCGCTTTCCAGTCCGGAAACCTGTCTGCGCACTGC	5205
Db	5141	GTGTATACTGGCTTAACTATGCGGCATCAGACAGATTGTACTGAGAGTGCACATATGC	5200
Qy	5206	ATTAAATGAATCGGCCAAACGCGCGGGAGAGCGGTTTGCATTGCGGCGCTCTTCGCTT	5265
Db	5201	GGTGTGAATACCGCAACAGATGCGTGAAGGAGAAAAATACCGCATCAGGCGCTCTTCGCTT	5260
Qy	5266	CCTCGCTCACTGACTCGCTCGCTCGGTTCGGTTCGGCTCGGCGAGCGGTATCAGCTCACT	5325
Db	5261	CCTCGCTCACTGACTGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACT	5320
Qy	5326	CAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAACATGTAG	5385
Db	5321	CAAGGCGGTAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAACATGTAG	5380
Qy	5386	CAAAAGCCAGCAAAAGGCCAGAAACGGTAAAGGCCGCGTTCGTGGCGTTTTCCTATA	5445
Db	5381	CAAAAGCCAGCAAAAGGCCAGAAACGGTAAAGGCCGCGTTCGTGGCGTTTTCCTATA	5440
Qy	5446	GGCTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGGAACCC	5505
Db	5441	GGCTCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGGAACCC	5500
Qy	5506	CGACAGACTATAAAGATACAGGCGGTTTCCCCCTGGAAAGTCCCTCGTCGCTCTCTCTG	5565
Db	5501	CGACAGACTATAAAGATACAGGCGGTTTCCCCCTGGAAAGTCCCTCGTCGCTCTCTCTG	5560
Qy	5566	TTCCGACCTTGCCTTTACCGGATACCTGTGCGCTTTTCTCCCTTCGGGAAGCGTGGCG	5625
Db	5561	TTCCGACCTTGCCTTTACCGGATACCTGTGCGCTTTTCTCCCTTCGGGAAGCGTGGCG	5620
Qy	5626	TTTCTCATAGCTCACCGTGTAGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAGCTCG	5685
Db	5621	TTTCTCATAGCTCACCGTGTAGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAGCTCG	5680
Qy	5686	GCTGTGTGACGAACCCCGCTTCAGCCGCAACGCTGCGCTTATCCGCTTAACATCTGTC	5745
Db	5681	GCTGTGTGACGAACCCCGCTTCAGCCGCAACGCTGCGCTTATCCGCTTAACATCTGTC	5740
Qy	5746	TTGAGTCCAAACCCGGTAAGACAGCTTATCGCCACTGGCAGCAGCACTGGTAAACAGA	5805

Db 5741 TTGAGTCCAAACCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA 5800
 QY 5806 TTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGCGCTTAACCTACG 5865
 Db 5801 TTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGCGCTTAACCTACG 5860
 QY 5866 GCTACACTAAGAGGACAGATTTTGGTATCTGGCTCTGCTGAAGCCAGTACCTTCGGAA 5925
 Db 5861 GCTACACTAAGAGGACAGATTTTGGTATCTGGCTCTGCTGAAGCCAGTACCTTCGGAA 5920
 QY 5926 AAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACCAACCGCTGTAGCGGTGTTTTTTG 5985
 Db 5921 AAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACCAACCGCTGTAGCGGTGTTTTTTG 5980
 QY 5986 TTTGCAAGCAGCAGATTACGCGCAGAAAAGGATCTCAAGAGATCCCTTTGATCTTTT 6045
 Db 5981 TTTGCAAGCAGCAGATTACGCGCAGAAAAGGATCTCAAGAGATCCCTTTGATCTTTT 6040
 QY 6046 CTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAATGAT 6105
 Db 6041 CTACGGGGTCTGACGCTCAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAATGAT 6100
 QY 6106 TATCAAAAAGGATCTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCT 6165
 Db 6101 TATCAAAAAGGATCTTCACTAGATCCCTTTTAAATTAATAATGAAGTTTAAATCAATCT 6160
 QY 6166 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTATCAGTGAGGCACCTA 6225
 Db 6161 AAGATATATAGTAAACCTTGGTCTGACAGTTACCAATGCTTATCAGTGAGGCACCTA 6220
 QY 6226 TCTCAGCGATCTGTCTATTTGTTTCCATCCATAGTTGCTTCCGCTGCTGTGATGATAA 6285
 Db 6221 TCTCAGCGATCTGTCTATTTGTTTCCATCCATAGTTGCTTCCGCTGCTGTGATGATAA 6280
 QY 6286 CTACAGTACGGAGGGCTTACCATCTGCCCCAGTCTGCAATGATACCGAGAGCCAC 6345
 Db 6281 CTACAGTACGGAGGGCTTACCATCTGCCCCAGTCTGCAATGATACCGAGAGCCAC 6340
 QY 6346 GTCCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCCAGAA 6405
 Db 6341 GTCCACCGCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGCCAGAA 6400
 QY 6406 GTGGTCTCTCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGCGGAGCTAGAG 6465
 Db 6401 GTGGTCTCTCAACTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGCGGAGCTAGAG 6460
 QY 6466 TAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTTGGCCATTTGTACAGGC-TCGTGG 6524
 Db 6461 TAAGTAGTTGCGCAGTTAATAGTTTGGCAACGTTTGGCCATTTGTACAGGCATCGTGG 6520
 QY 6525 TGTACGCTCGTGGTTGTATGGCTTCAATCAGCTCCGCTCCCAACCATCAAGCGAG 6584
 Db 6521 TGTACGCTCGTGGTTGTATGGCTTCAATCAGCTCCGCTCCCAACCATCAAGCGAG 6580
 QY 6585 TTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCTCCGATCGTTG 6644
 Db 6581 TTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCCTTGGTCTCTCCGATCGTTG 6640
 QY 6645 TCAGAAAGTAAAGTTGGCGCAGTTTATCACTCATGTTTATGGCAGCACTGCATAATTC 6704
 Db 6641 TCAGAAAGTAAAGTTGGCGCAGTTTATCACTCATGTTTATGGCAGCACTGCATAATTC 6700
 QY 6705 TTACTGTATGCAATCCGTAAGTATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCAT 6764
 Db 6701 TTACTGTATGCAATCCGTAAGTATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCAT 6760
 QY 6765 TCTGAGAAAGTGTATGCGCGCACCGAGTTGCTTCTTGGCCGCGTCAATACCGGATAA 6824
 Db 6761 TCTGAGAAAGTGTATGCGCGCACCGAGTTGCTTCTTGGCCGCGTCAACCGGATAA 6820
 QY 6825 CCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGTCTTTCGGGCGAA 6884
 Db 6821 CCGCGCCACATAGCAGAACTTTAAAGTGTCTATCATTTGAAAAAGTCTTTCGGGCGAA 6880

QY 6885 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCA 6944
 Db 6881 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCA 6940
 QY 6945 ACTGATCTTCAAGCATCTTTTACTTTTACAGCAGGTTTCTGGGTGAGCAAAAACAGGAAGC 7004
 Db 6941 ACTGATCTTCAAGCATCTTTTACTTTTACAGCAGGTTTCTGGGTGAGCAAAAACAGGAAGC 7000
 QY 7005 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 7064
 Db 7001 AAAATGCCGCAAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATACTCTTCC 7060
 QY 7065 TTTTTCATATTTATTTGAAGCATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATTG 7124
 Db 7061 TTTTTCATATTTATTTGAAGCATTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATTG 7120
 QY 7125 AATGATTTTAAAAAATAAACAATAGGGTTTCCGCGACATTTCCCGAAAAAGTGCCAC 7184
 Db 7121 AATGATTTTAAAAAATAAACAATAGGGTTTCCGCGACATTTCCCGAAAAAGTGCCAC 7180
 QY 7185 CTGAGCTCTAAGAAACCATTTATTCATGACATTAACCTATATAAATAGGCGTATCACGA 7244
 Db 7181 CTGAGCTCTAAGAAACCATTTATTCATGACATTAACCTATATAAATAGGCGTATCACGA 7240
 QY 7245 GGCCCTTTTCTGCT 7257
 Db 7241 GGCCCTTTTCTGCT 7253
 RESULT 11
 AB086386
 LOCUS AB086386 6076 bp DNA circular SYN 14-NOV-2003
 DEFINITION Retroviral vector pCX4pur DNA, complete sequence.
 ACCESSION AB086386
 VERSION AB086386.1 GI:26449194
 KEYWORDS
 SOURCE Retroviral vector pCX4pur
 ORGANISM Retroviral vector pCX4pur
 REFERENCE 1
 AUTHORS Akagi,T., Sasai,K. and Hanafusa,H.
 TITLE Refractory nature of normal human diploid fibroblasts with respect to oncogene-mediated transformation
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (23), 13567-13572 (2003)
 MEDLINE 22975258
 PUBMED 14597713
 REFERENCE 2 (bases 1 to 6076)
 AUTHORS Akagi,T.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUN-2002) Tsuyoshi Akagi, Osaka Bioscience Institute, Molecular Oncology; Furuedai 6-2-4, Suita, Osaka 565-0874, Japan (E-mail: takagi@obi.or.jp; Tel:81-6-6872-4834, Fax:81-6-6871-7521)
 Location/Qualifiers
 FEATURES
 source
 1..6076
 /organism="Retroviral vector pCX4pur"
 /mol_type="other DNA"
 /db_xref="taxon:197852"
 enhancer
 1..466
 /note="CMV IE enhancer"
 LTR
 467..762
 /note="R-U5 region of MolV 5' LTR"
 misc_signal
 763..1667
 /note="psi+; extended packaging signal"
 misc_feature
 1668..2045
 /note="Mo-MuLV pol region containing splice acceptor"
 misc_feature
 2044..2106
 /note="multiple cloning site"
 RBS
 2107..2716
 /note="IRES sequence from EMCV"
 gene
 2718..3317
 /gene="pur"
 CDS
 2718..3317

LTR
gene
CDS

```
/gene="pur"  
/codon_start=1  
/product="puromycin-N-acetyl transferase"  
/protein_id="BAC41726.1"  
/db_xref="GI:26449195"  
/translation="MTBYKPTVRLATRDVPRVATLAAAFADYPATRHVTDPDRH1E  
RVTELOELFLTRVLGDLGKVVADGAAVA VMTTPESVEAGVPAEIGPMAELSGSR  
LAAQQOEGLLAPHRPEKPAWPLATVGVSPHQKGLGSVVLPGVAAERAGVPAFL  
ETSAPNLNPFYERLGFVIADVECPKORATWCMRKPGA"  
3371..3962  
/note="3' LTR of Mo-MuLV"  
complement(5015..5875)  
/gene="bla"  
complement(5015..5875)  
/gene="bla"  
/codon_start=1  
/product="beta-lactamase"  
/protein_id="BAC41727.1"  
/db_xref="GI:26449196"  
/translation="MSIQHFVALIPFFAFLPFAHPETLVKVKDAEDQLGARVGY  
IELDLSGKILLESFRRPELSPMFTSMVTEFVKLLGAVLSRVDAEQEOLGRRIHYSONDLVE  
YSPVTEKHLDTVRBELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGDHVRLL  
DRWPELNEALPNDERDTMEVAMATTLRLKLLTGELLTLASROQLIDWMEADKVGPL  
LRSALPAGWFTADKSGAGERSGI IAAALGPDGKPSRIIVVIYTTGSAATWDERNRQIA  
EIGASLIKHW"
```

ORIGIN

Query Match 40.7%; Score 3113; DB 12; Length 6076;
Best Local Similarity 73.3%; Pred. No. 0;
Matches 4708; Conservative 0; Mismatches 895; Indels 818; Gaps 24;

Qy	843	TTCTAGAGAACCATCAGATGTTTCCAGGTGCCCAAGACCTGAATGACCCCTGTGCCT	902
Db	466	TGCTAGAGAACCATCAGATGTTTCCAGGTGCCCAAGACCTGAATGACCCCTGTGCCT	525
Qy	903	TATTTGAACCAATCAGTTCCTCTCGCTTCTGTTTCGGCGCTTCTGCTCCCGAG	962
Db	526	TATTTGAACCAATCAGTTCCTCTCGCTTCTGTTTCGGCGCTTCTGCTCCCGAG	585
Qy	963	CTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGCGTCG	1022
Db	586	CTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATGACTGAGTCG	645
Qy	1023	CCCGGGTACCCGTAATCCCAATAAGCCTTTGCTGTTTGCATCCGAATCGTGACTCG	1082
Db	646	CCCGGGTACCCGTAATCCCAATAAGCCTTTGCTGTTTGCATCCGAATCGTGACTCG	704
Qy	1083	TGATCCTTGGAGAGGTCTCTCAGATTGATGACTGCCC-ACCTCGGGGGTCTTTCAATT	1141
Db	705	TGTTCTTGGAGAGGTCTCTCTGAGTGATTGACTACCCCGTACGCGGGGTCTTTCAATT	764
Qy	1142	GGAGGTTCCACCGAGATTTGGAGACCCCTGCGCGCGCAGTCTCCGATAGACTGCGTCG	1201
Db	765	GGGGGCTCGTCCGGATCGGGAGACCCCTGCGCGCGCAGTCTCCGATAGACTGCGTCG	823
Qy	1202	GTAAGCTGGCCAGCAACTATCTGTGTCTGTCGGAITGTCTAGTGTCTATGACTGATTTT	1261
Db	824	GTAAGCTGGCCAGCAACTATCTGTGTCTGTCGGAITGTCTAGTGTCTATGACTGATTTT	883
Qy	1262	ATGGCCCTGGTTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGA	1321
Db	884	ATGGCCCTGGTTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGA	943
Qy	1322	ACTGACGAGTTCGGAAACACCGCGCGCAACCTCGGGAGAGTCCACGGACTTCGGGGG	1381
Db	944	ACTGACGAGTTCGGAAACACCGCGCGCAACCTCGGGAGAGTCCACGGACTTCGGGGG	1003
Qy	1382	CGTTTTTGTGCCCGACTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTTGTGCACC	1441
Db	1004	CGTTTTTGTGCCCGACTGAGTCCAAATCCCGATCGTTTGGACTCTTTTGTGCACC	1063
Qy	1442	CCCTTAGAGAGGAGATATGTGTTCTGTGTTAGGAGACGAGAACCTAAACAGTTTCCGCC	1501

Db	1064	CCCCTAATAGAGGGATATGTGGTCTCTGGTAGGAGACGAGAACCTAAACAGTTCCCGCC	1123
Qy	1502	TCCGCTGTAATTTTGTCTTTTCGGTTTGGAGCCGAGCCGCGCGCTCTTGTCTGCTG	1561
Db	1124	TCCGCTGTAATTTTGTCTTTTCGGTTTGGAGCCGAGCCGCGCGCTCTTGTCTGCTG	1183
Qy	1562	CAGCATCGTTCGTGTGTCTCTCTCTGCTGACTGTGTCTGTATTTGTCTGAAATATGGG	1621
Db	1184	CAGCATCGTTCGTGTGTCTCTCTCTGCTGACTGTGTCTGTATTTGTCTGAAATATGGG	1243
Qy	1622	CCCGGGCTAGACTGTATACCACTCCCTTAAGTTTTCACCTTAGGTCACTGGAAGATGTGGA	1681
Db	1244	CC-----AGACTGTTTACCCTCCCTTAAGTTTTCACCTTAGGTCACTGGAAGATGTGGA	1297
Qy	1682	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAAGAGACGTTGGGTACTCTTGTCTC	1741
Db	1298	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAAGAGACGTTGGGTACTCTTGTCTC	1357
Qy	1742	TGCAGATGGCCAAACCTTTAAACGTGGATGGCCGCGGAGACGGACCTTTAAACCGAGACT	1801
Db	1358	TGCAGATGGCCAAACCTTTAAACGTGGATGGCCGCGGAGACGGACCTTTAAACCGAGACT	1417
Qy	1802	CATCACCCAGGTTAAGATCAAGGTCTTTTTCACCTGGCCCGCATGAGACCCAGACCCAGGT	1861
Db	1418	CATCACCCAGGTTAAGATCAAGGTCTTTTTCACCTGGCCCGCATGAGACCCAGACCCAGGT	1477
Qy	1862	CCCCTACATCGTACCTGGGAAGCCTTGGCTTTTTCAGCCCGCCCTCTCTGGGTCAAGCCCTT	1921
Db	1478	CCCCTACATCGTACCTGGGAAGCCTTGGCTTTTTCAGCCCGCCCTCTCTGGGTCAAGCCCTT	1537
Qy	1922	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCATTCGCGCCCGCTCTCTCCCTTTGAACC	1981
Db	1538	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCATTCGCGCCCGCTCTCTCCCTTTGAACC	1597
Qy	1982	TCCTCGTTTCGACCCCGCTCGATCCTCCCTTATTCAGAGCCCTCACTCTCTCTAGGGCG	2041
Db	1598	TCCTCGTTTCGACCCCGCTCGATCCTCCCTTATTCAGAGCCCTCACTCTCTCTAGGGCG	1657
Qy	2042	CCCCTATGCGCATATGAGATCTTATATGGGGACCCCGCCCTTGTAAACTTCCCTGA	2101
Db	1658	CCCCTATGCGCATATGAGATCTTATATGGGGACCCCGCCCTTGTAAACTTCCCTGA	1717
Qy	2102	CCCTGACATGACAGAGTTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCT-TCTA	2160
Db	1718	CCCTGACTTGACAAGAGTTACTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCTCTCTA	1777
Qy	2161	CTTAGTCCAGACGGAAGTCTGAGACCTCTGGCGGAGCCTTACCAAGAACCACTGGAACG	2220
Db	1778	CTTAGTCCAGACGGAAGTCTGAGACCTCTGGCGGAGCCTTACCAAGAACCACTGGAACG	1837
Qy	2221	ACCGGTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGACACACAGAC	2280
Db	1838	ACCGGTGGTACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGACACACAGAC	1897
Qy	2281	TAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACACCCACCGCC	2340
Db	1898	TAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACACCCACCGCC	1957
Qy	2341	CCTCAAGTAGAGCGGATCGGATACAGCCCGCCGACGCTGAAGGCTGCGGACCC	2400
Db	1958	CCTCAAGTAGAGCGGATCGGATACAGCCCGCCGACGCTGAAGGCTGCGGACCC	2017
Qy	2401	CGGGGTGGACCACTCTCTAGACTGCCATGGAGTGGAGTGTATCATCTCTTCTTGTA	2460
Db	2018	CGGGGTGGACCACTCTCTAGACTGCCGATGCCGATGCCAGTGTGGTGTAGCGGAATTCGGG	2075
Qy	2461	GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGC	2520
Db	2076	---ACCGACCGGTGGCGCCGGGTAACTTGTGTAACTTAAGCTAGTACCAAGGTTTC	2132
Qy	2521	GCCAGCGTGGGTGACAGGTGACATCATCTGTAAAGCCAGTCAAGGATGGGTACTTCT	2580
Db	2133	CCTCTAGCGGGATCAATTCGCGCCCGCCCGCTTAACTTACT-----	2174

QY 2581 GTAGCTTGTGACAGCAGAGCCAGGTAAGGCTCCAAAGCTGCTGATCTACTTGGACATCC 2640
 Db 2175 ---GGCCGAAGCGCTTGGAAATAGGCCGCTGTGGTTTGTCTATATGTATATTTCCACC 2231
 QY 2641 ACCCGCACACTGGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACC 2700
 Db 2232 ATATTGCCCTCTTTTG------GCAATGTAGGGCCCGGAACCTGGCCCTGCTCTCTTTG 2285
 QY 2701 TTCACCATCAGCAGCTTCAGCCAGAGGACATCGCCACCTACTACTGCGCAGCAATATAGC 2760
 Db 2286 ACAGCATTCCTAGGGGCTTTTCCCTCTCGCAAGAGGAATCAAGGCTGTGTGAATGTC 2345
 QY 2761 CTCTATCGGTCTGTGGCCCAAGGACCAAGGTGGAATCAAAACGAGGTGGCTCAGATCG 2820
 Db 2346 GTGA--GGAAGAGTTCCTCTGGAAGCTTCTTGAAGACAAACACGCTGTAGCGACCC 2403
 QY 2821 GGTGATCCGGCTCTGTGGCTCAGGATCGAGGTCCAACTGGTGGAGCGGTGGAGGT 2880
 Db 2404 TTTGAGGCGAGGAACCCCGACCTGGCGACAGGTGCTCTGGCGCAAAAGCCACGTG 2463
 QY 2881 GTTGTGCACTGGCGGTCCCTGGCTGTCTGCTCGGATCTGGCTTCGATTTTACC 2940
 Db 2464 TATAAGATACCTCGCAAGGCGGCACAAACCCAGTGCACGTTGTAGTTGGATGTTG 2523
 QY 2941 ACATATTGATGATTTGGGTGAGACAGGCACCTGGAAAGGTCCTTGAAGTGGAGAA 3000
 Db 2524 TGAAGAAGTCAATGGCTCTCTCTCAGCGTATTCMAAGGGGTGAAGATGCCAGA 2583
 QY 3001 ATTTCATCCAGATGACGATGATTAATCTATGGCGCTCTCTAAAGGATAGATTTACAATA 3060
 Db 2584 AGGTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCACATGCTTTTACATGTGTT 2643
 QY 3061 TCGGAGACAGCCCAAGAACACATTTGCTGCAATGGACAGCTGAGACCCGAGAC 3120
 Db 2644 AGTCGAGGTTAAAAACGTCCTA-----GGCCCCCGCAACACCGG 2683
 QY 3121 ACCGGGCTCTATTTTGTGCAAGCTTTTACTTTCGGCTTCCCTGTTTCTTATTTGGGC 3180
 Db 2684 GACGTGTTTCTCTTTGAAAAACACGATTAATACATGACCGAGTACAGCCACG- --- 2739
 QY 3181 CAAGGACCCCGTCAACCGTCTCCAGTGTAAAGCCCAACACAGCGCAGCGCGGACCA 3240
 Db 2740 -----TGCGCTCGCCACCCGCGACGAGCTCCCGCGCG 2775
 QY 3241 CCAAACCCGCGCCACATCGGTGCGAGCCCTGTCTCGGCGCCAGAGCGGCTCGG 3300
 Db 2776 TAGCGACCTCGCGCGGCTTTCGCGACTACCCCGCCACGCGCCACACCGTCGACCGG 2835
 QY 3301 CCAGCGCGGGGGCGAGTGACACAGAGGGGCTGACTTCGCTCGGATCCCAACTC 3360
 Db 2836 ACCGCCACATCGAGCG-----GGTCACGAGCTGCAAGAACTCTTCTCAGCGCGTGGG 2891
 QY 3361 TGCTACTGCTGATGGAATCTCTTCTATATGTGTCTATCTCACTGCTTGTTCCTG 3420
 Db 2892 CTCGACATCGCAAGGTGTGGGTGCGGACGAGCGCGCGCTGCGGTCTGGACCAAG 2951
 QY 3421 AGAGTGAATTGACAGGAGCGCAGACCCCGCGGTAACAGAGGGCCAGAACAGCTC 3480
 Db 2952 CCGGAGAGGCTCGAAGCGGGCGGTGTTTCGCGGAGATCGGCGCGCATGCGCGAGTTG 3011
 QY 3481 TATAACGAGCTCAATCTAGACGAGAGAGGATGATGATTTTGGACAAGAGAGCTGGC 3540
 Db 3012 AGCGGTTCCCGGTCGCGCGAGCAACAGATGGAAGGCTCTCTGGCGCGCACCGGCC 3071
 QY 3541 CGGGACCTCTGATGCGGGGAAAGCCGAGAGGAAGAACCTCTCAGGAAGGCTCTGACAT 3600
 Db 3072 AAGGACCCCGGTGTT-----CCTGGCCACCGTCGGGCTCTC 3109
 QY 3601 GAATCTCAGAAAGATTAAGATGGCGAGGCTCAGTGAGATTTGGATGAAGCGAGCGC 3660
 Db 3110 GCCCGACCAACAGGCAAGGCTCTGGCGAGCGCGCTCGTCTCCCGGAGTGGAGCGCGC 3169

QY 3661 CGAGGGGCAAGGGGCACGATGGCCTTTTACAGGGTCTCAGTACAGCCCAAGACACACC 3720
 Db 3170 CGAGCGCGCGGG-----TGCCGCGCTTCTCTGGAGACC 3203
 QY 3721 TAGCAGCGCTTTCATGTCAGGCGCTGCCCTCGCTAACTTCGACGCGCGCGGATCCG 3780
 Db 3204 TCCGCGCGCGCAACCTCCCTTCTACGAGCGGCTCGGCTTTCACCGTCAACCGCCGCGTC 3263
 QY 3781 GATTAGTCCAAATTTGTAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAAC 3840
 Db 3264 GAGTCCCGAAGGACCGCGGACCTGTGTGATGACCCGAAAGCCGCTGCTGAGCGCTC 3323
 QY 3841 AATATCACCAAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTTA 3900
 Db 3324 GACA-----CGATAAAATAAAGATTTTATTTA 3351
 QY 3901 GTCTCCAGAAAAGGGGGGAATGAAGAACCCTGTAAGTTTGGCAAGCTAGCTTAAG 3960
 Db 3352 GTCTCCAGAAAAGGGGGGAATGAAGAACCCTGTAAGTTTGGCAAGCTAGCTTAAG 3411
 QY 3961 TAAAGCCATTTTGCAGGCGATGG-AAAATACATACATCAGATAGAGAGTTCAGATCAA 4019
 Db 3412 TAAGCCCATTTTGAAGGCGATGGAATAATACATACTGAGATAGAGAGTTCAGATCAA 3471
 QY 4020 GGTAGGAAACAGA-GAGACAGCAGATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 4078
 Db 3472 GGTAGGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGTT 3531
 QY 4079 CCGTGGCC- GGTAGGCGCAAGAACAGTTGGAACAGGAGATATGGGCCAAACAGGATAT 4137
 Db 3532 CCGTGGCGGCTCAGGCGCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 3591
 QY 4138 CTGTGTAAGCAGTTCTGCGCGCTCAGGCGCAAGAACAGATGTCCTCCAGAGTCCGT 4197
 Db 3592 CTGTGTAAGCAGTTCTGCGCGGCTCAGGCGCAAGAACAGATGTCCTCCAGAGTCCGT 3651
 QY 4198 CCGCGCTCAGCAGTTCTTAGAAGACCATCAGATGTTTCCAGGGTGCCTCCAGAGCTTGA 4257
 Db 3652 CCAGCGCTCAGCAGTTCTTAGAAGACCATCAGATGTTTCCAGGGTGCCTCCAGAGCTTGA 3711
 QY 4258 AATGACCTGTGCTTATTTGAATCAACATCAGTTCGTTCTCGTTCTGTTCCGCGG 4317
 Db 3712 AATGACCTGTGCTTATTTGAATCAACATCAGTTCGTTCTCGTTCTGTTCCGCGG 3771
 QY 4318 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTC 4377
 Db 3772 CTTCGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCAGTCTC 3831
 QY 4378 CGATAGACTGCGTCCCGGGTACCGGTCTCTCAATAAACCTCTTGCAGTTGCATCG 4437
 Db 3832 CGATTGACTGAGTCCCGGGGTACCCGTGTATCCAAATAAACCTCTTGCAGTTGCATCG 3891
 QY 4438 ACTCGTGTCTGCTGTTCTTGGAGGGTCT-CTCTGAGTGAATTGACTACCGCTCAGG 4496
 Db 3892 ACTTGTGCTCTGCTGTTCTTGGAGGGTCTCTCTCTGAGTGAATTGACTACCGCTCAGG 3951
 QY 4497 GGTCTTTCAGTTTCTCCCACTACAGGTTCTCACTAACAATTCCTGATGTCGCGAGG 4556
 Db 3952 GGG----- 3954
 QY 4557 ACTCGGTGAGCGCGGTTTGTGTTTATATAAATAAAGAACAGTGTCTCCCTTCAAGCCA 4616
 Db 3955 ----- 3954
 QY 4617 GACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTATCTGC 4676
 Db 3955 ----- 3954
 QY 4677 CACACGACTTTTAAAGATTTTATGCTCTCGATGAGGGATTTAGTCAATCTATCTCTCG 4736
 Db 3955 -----GTCTTTCATGCGAGCTGATATCAAAATTAATTTGG 3990
 QY 4737 TCTATTTGCTGGCTTCTCCGCTATTTTAAATTTTCTAGTTTGCATCTCCCTTCTCTGAGACA 4796

Db	3991	 TTTTTTT	-----	3999
Qy	4797	CGCGCATTCAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAGGTTGCCTGGGC	4856	
Db	3999	----- CTTAAGTATTACATTAAATGGCC	4022	
Qy	4857	TCAGTGTGAGATTTTGGCCATAAAAGGGGTCTGCCCCCTGTGTACAGACAGATCGGAATC	4916	
Db	4023	ATAGT-----	4027	
Qy	4917	TAGAGTGCATATCTCAGAGTCCCCGCGTTCCGGGGCTCTGATCTCAGGGCATCTTTGCCT	4976	
Db	4028	-----	4027	
Qy	4977	AGATATCCTCTACGCGCGAGCATGCTGTGGCGGGTACCGAGCTCGAATTCGTAATCATGG	5036	
Db	4028	-----	4027	
Qy	5037	TCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCAAAATCCACACACATACGAGCC	5096	
Db	4028	-----	4027	
Qy	5097	GGAAGCATTAAGTCTAAAGCCTGGGTGCCTAATGAGTGAGCTAACTCACTAATTGCG	5156	
Db	4028	-----	4027	
Qy	5157	TTGCGCTCACTGCCCGCTTTCAGTCCGGAAACCTGTGTCGCCAGCTGCATTAAATGAATC	5216	
Db	4028	----- TGCATTAAATGAATC	4041	
Qy	5217	GGCAACGCGGGGAGAGCGGTTTGCGTATTTGGCGCTCTTCCGCTTCTCCGCTCACT	5276	
Db	4042	GGCAACGCGCGGGAGAGCGGTTTGCGTATTTGGCGCTCTTCCGCTTCTCCGCTCACT	4101	
Qy	5277	GACTCGCTGGCTCGGTCGTTCCGCTCGCGGAGCGGTATCAGTCACTCAAAAGGCGGTA	5336	
Db	4102	GACTCGCTGGCTCGGTCGTTCCGCTCGCGGAGCGGTATCAGTCACTCAAAAGGCGGTA	4161	
Qy	5337	ATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAAAGGCCAG	5396	
Db	4162	ATACGGTTATCCACAGAAATCAGGGGATAACGCGAGGAAGAACATGTGAGCAAAAGGCCAG	4221	
Qy	5397	CAAAAGGCCAGGAACCGGTAAAGAGCCGCGTTGTGCGCGTTTTTCCATAGGCTCCGCCCC	5456	
Db	4222	CAAAAGGCCAGGAACCGGTAAAGAGCCGCGTTGTGCGCGTTTTTCCATAGGCTCCGCCCC	4281	
Qy	5457	CCTGACAGCATCAAAAAATTCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTA	5516	
Db	4282	CCTGACAGCATCAAAAAATTCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGACTA	4341	
Qy	5517	TAAAGATACCAGGCGTTTCCCGCTGGAAGCTCCCTCGTGCGCTCTCTGTTCGACCCCTG	5576	
Db	4342	TAAAGATACCAGGCGTTTCCCGCTGGAAGCTCCCGTGCGCTCTCTGTTCGACCCCTG	4401	
Qy	5577	CCGCTTACCAGATACCTGTCCGCGTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGC	5636	
Db	4402	CCGCTTACCAGATACCTGTCCGCGTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGC	4461	
Qy	5637	TCAGCTGTAGGTATCTCAGTTCCGTTGTAAGTTCGTTCCGCTCCAAGCTGGGCTGTGTGCAC	5696	
Db	4462	TCAGCTGTAGGTATCTCAGTTCCGTTGTAAGTTCGTTCCGCTCCAAGCTGGGCTGTGTGCAC	4521	
Qy	5697	GAACCCCGGTTTCAGCCCGACCGCTCGCCTTATCCGGTAACTTATCGTCTTGAGTCCAAC	5756	
Db	4522	GAACCCCGGTTTCAGCCCGACCGCTCGCCTTATCCGGTAACTTATCGTCTTGAGTCCAAC	4581	
Qy	5757	CCGTTAAGACACGACTTATCCCACTCGGACAGCCACTCGTAAACAGGATTACGACGCG	5816	
Db	4582	CCGTTAAGACACGACTTATCCCACTCGGACAGCCACTCGTAAACAGGATTACGACGCG	4641	
Qy	5817	AGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGCTACACTAGA	5876	

4642	AGGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGCGCTACACTAGA	4701
5877	AGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGAGTTACCTTCGGAAAAAGAGTTGGT	5936
4702	AGGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGT	4761
5937	AGCTCTTTGATCCGGCAAAACAAACGACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAG	5996
4762	AGCTCTTTGATCCGGCAAAACAAACGACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAG	4821
5997	CAGATTACGGCGAGAAAAAAGGATCTCAAGAGAGATCCCTTTTGATCTTTTCTACGGGCTCT	6056
4822	CAGATTACGGCGAGAAAAAAGGATCTCAAGAGATCCCTTTTGATCTTTTCTACGGGCTCT	4881
6057	GACGCTCAGTGGAAACGAAAACTCACGCTTAAAGGATTTTGGTCAATGAGATTTATCAAAAAGG	6116
4882	GACGCTCAGTGGAAACGAAAACTCACGCTTAAAGGATTTTGGTCAATGAGATTTATCAAAAAGG	4941
6117	ATCTTACCTAGATCTCTTTTAAATTTAAATGAAGTTTTAAATCAATCTTAAAGTATATAT	6176
4942	ATCTTACCTAGATCTCTTTT-----GCGCGCGCGCGCAAACTCAATCTAAAGTATATAT	4994
6177	GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACCTATCTCAGCGATC	6236
4995	GAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACCTATCTCAGCGATC	5054
6237	TGTCATATTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGG	6296
5055	TGTCATATTCGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGG	5114
6297	GAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGCT	6356
5115	GAGGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAGCTCACCGGCT	5174
6357	CCAGATTATACGCAATAAACACGACGACGCGGAAGCGCCGAGCGCAGAGTGTGCTGCA	6416
5175	CCAGATTATACGCAONATAAACACGACGACGCGGAAGCGCCGAGCGCAGAGTGTGCTGCA	5234
6417	ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAGAGTAAGTAGTTCG	6476
5235	ACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAGAGTAAGTAGTTCG	5294
6477	CCAGTTAATAGTTTGGCCAAAGTTTGTTCGCAATTTGCTACAGGC-TCGTFGTGTCAACGCTCG	6535
5295	CCAGTTAATAGTTTGGCCAAAGTTTGTTCGCAATTTGCTACAGGCATCTGTGTGTCAACGCTCG	5354
6536	TCGTTTGGTATGGTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC	6595
5355	TCGTTTGGTATGGTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC	5414
6596	CCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCGATCGTTGTGCAAGAGTAAG	6655
5415	CCCATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCGATCGTTGTGCAAGAGTAAG	5474
6656	TTGSCGCGAGTTATCACTCATGTTTATGGCAGCACTGCATAAATTTCTTTACTGTCATG	6715
5475	TTGSCGCGAGTTATCACTCATGTTTATGGCAGCACTGCATAAATTTCTTTACTGTCATG	5534
6716	CCATCCGTAAGATGCTTTTCTGTACTGTGTAGTACTCAACGAAAGTCATTTCTGAGAAATAG	6775
5535	CCATCCGTAAGATGCTTTTCTGTACTGTGTAGTACTCAACGAAAGTCATTTCTGAGAAATAG	5594
6776	TGTAATGGCGGACCGAGTTGCTCTCTCCGCGCTCAATACGGGATAATACCGCGCCACAT	6835
5595	TGTAATGGCGGACCGAGTTGCTCTCTCCGCGCGCTCAACGCGGATAATACCGCGCCACAT	5654
6836	AGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTCTTCGGGGCGGAAAACTCTCAAGG	6895
5655	AGCAGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTTCTTCGGGGCGGAAAACTCTCAAGG	5714
6896	ATCTTACCGCTGTGTAGATCCAGTTGATGTAAACCACTCGTGTGCAACCAACTGATCTTCA	6955
5715	ATCTTACCGCTGTGTAGATCCAGTTGATGTAAACCACTCGTGTGCAACCAACTGATCTTCA	5774

QY 5223 CGCGGGGAGAGGGGTTTGGGTATTTGGGGCGCTCTTCGGCTTCTCGCTCACTGACTCG 5282
Db 3094 CGCGGGGAGAGGGGTTTGGGTATTTGGGGCGCTCTTCGGCTTCTCGCTCACTGACTCG 3153
QY 5283 CTGCGCTCGGTCTGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGGGGTAATACGG 5342
Db 3154 CTGCGCTCGGTCTGGCTCGCGGAGCGGTATCAGCTCACTCAAAAGGGGTAATACGG 3213
QY 5343 TTATCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG 5402
Db 3214 TTATCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG 3273
QY 5403 GCCAGGAACGTAATAAAGGCCGCTTGTCTGGCTTTTTCATAGGCTCGGCCCCCTGAC 5462
Db 3274 GCCAGGAACGTAATAAAGGCCGCTTGTCTGGCTTTTTCATAGGCTCGGCCCCCTGAC 3333
QY 5463 GAGCATCACAAAATCAGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAGA 5522
Db 3334 GAGCATCACAAAATCAGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAGA 3393
QY 5523 TACCAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTGTTCGACCCCTCGCCTT 5582
Db 3394 TACCAGGCTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTGTTCGACCCCTCGCCTT 3453
QY 5583 ACCGATACCTGTCGGCTTTCTCGCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGC 5642
Db 3454 ACCGATACCTGTCGGCTTTCTCGCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGC 3513
QY 5643 TGTAAGTATCTCAGTTCGGTGTAGTTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 5702
Db 3514 TGTAAGTATCTCAGTTCGGTGTAGTTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 3573
QY 5703 CCCGTTTCCGCGGACCGCTGCGCTTATCCGCTTAATCCGCTTAATCCGCTTAATCCGCTTA 5762
Db 3574 CCCGTTTCCGCGGACCGCTGCGCTTATCCGCTTAATCCGCTTAATCCGCTTAATCCGCTTA 3633
QY 5763 AGACAGACTTATCGCCACTGGAGCGCTGCGCTTATCCGCTTAATCCGCTTAATCCGCTTA 5822
Db 3634 AGACAGACTTATCGCCACTGGAGCGCTGCGCTTATCCGCTTAATCCGCTTAATCCGCTTA 3693
QY 5823 GTAGCGGTGCTACAGAGTTCCTCAAGTGGTGGCTTAACCTACGGCTACCTAGAGGACA 5882
Db 3694 GTAGCGGTGCTACAGAGTTCCTCAAGTGGTGGCTTAACCTACGGCTACCTAGAGGACA 3753
QY 5883 GTATTTGGTATCTCGCTCTGCTGAAGCCAGTACCTTCGAAAAAGAGTTGGTAGCTCT 5942
Db 3754 GTATTTGGTATCTCGCTCTGCTGAAGCCAGTACCTTCGAAAAAGAGTTGGTAGCTCT 3813
QY 5943 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAAGCAGAGATT 6002
Db 3814 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAAGCAGAGATT 3873
QY 6003 ACGCGAGAAAAAGAGTCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTGAGCT 6062
Db 3874 ACGCGAGAAAAAGAGTCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTGAGCT 3933
QY 6063 CAGTGGAAACGAAAACTCAGCTTAAGGAGTTTGGTCAATGAGATTATCAAAAAGGATCTTC 6122
Db 3934 CAGTGGAAACGAAAACTCAGCTTAAGGAGTTTGGTCAATGAGATTATCAAAAAGGATCTTC 3993
QY 6123 ACCTAGATCTCTTTTAAATATAAATGAAGTTTAAATCAATCTAAAGATATATAGTAA 6182
Db 3994 ACCTAGATCTCTTTTAAATATAAATGAAGTTTAAATCAATCTAAAGATATATAGTAA 4053
QY 6183 ACTTGGTCTGACAGTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTA 6242
Db 4054 ACTTGGTCTGACAGTACCAATGCTTAATCAGTAGGACCTATCTCAGCGATCTGTCTA 4113
QY 6243 TTTGCTTCACTCATAGTTGCTGACTCCCGCTCGGTAGATAAATCAGATACGGAGGGC 6302
Db 4114 TTTGCTTCACTCATAGTTGCTGACTCCCGCTCGGTAGATAAATCAGATACGGAGGGC 4173
QY 6303 TTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGAT 6362

Db 4174 TTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGAT 4233
QY 6363 TTATCAGCAATAAACACGAGCCAGCGGAAAGGCGGAGCGCAGAAAGTGGTCTCGCAACTTTA 6422
Db 4234 TTATCAGCAATAAACACGAGCCAGCGGAAAGGCGGAGCGCAGAAAGTGGTCTCGCAACTTTA 4293
QY 6423 TCCGCTCCATCCAGTCTATAATTTGTTCGGGAAAGTACAGTAAGTAGTTCCGCAAGTT 6482
Db 4294 TCCGCTCCATCCAGTCTATAATTTGTTCGGGAAAGTACAGTAAGTAGTTCCGCAAGTT 4353
QY 6483 AATAGTTTTCGCAACGTTGTTGCCATTTCTACAGGC-TCCGTGGTGTCAAGCTCCGCTCGTTT 6541
Db 4354 AATAGTTTTCGCAACGTTGTTGCCATTTCTACAGGCATCGTGGTGTCAAGCTCCGCTCGTTT 4413
QY 6542 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG 6601
Db 4414 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG 4473
QY 6602 TTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTTCGATCGTTGTGAGAGTAAGTTGGCC 6661
Db 4474 TTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTTCGATCGTTGTGAGAGTAAGTTGGCC 4533
QY 6662 GCAGTGTATCACTCATGTTATGGCAGCACTGCATTAATTTCTTACTGTCTATGCCATCC 6721
Db 4534 GCAGTGTATCACTCATGTTATGGCAGCACTGCATTAATTTCTTACTGTCTATGCCATCC 4593
QY 6722 GTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTCGAATAAGTGTATG 6781
Db 4594 GTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTCGAATAAGTGTATG 4653
QY 6782 CGCGCAGCAGGTTGCTCTTCGCGCGCTCAATACGGGATAATACCGGCCACATAGCAGA 6841
Db 4654 CGCGCAGCAGGTTGCTCTTCGCGCGCTCAATACGGGATAATACCGGCCACATAGCAGA 4713
QY 6842 ACTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGGGGCGAAACTCTCAAGGATCTTA 6901
Db 4714 ACTTTAAAGTGTCTCATCTTGGAAACGTTCTTCGGGGCGAAACTCTCAAGGATCTTA 4773
QY 6902 CCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGACCCCACTGATCTTCAGCATCT 6961
Db 4774 CCGCTGTTGAGATCCAGTTCGATGTAACCACTCGTGACCCCACTGATCTTCAGCATCT 4833
QY 6962 TTATCTTTACACGCGTTCTGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAG 7021
Db 4834 TTATCTTTACACGCGTTCTGGTGAGCAAAAACAGAAAGGCAAAATGCCGCAAAAAG 4893
QY 7022 GGAATAAGGGCGACACGGAATGTTGAAATGTAATCTCTCTCTTTTCAATATATGA 7081
Db 4894 GGAATAAGGGCGACACGGAATGTTGAAATGTAATCTCTCTCTTTTCAATATATGA 4953
QY 7082 AGCATTTTATCAGGGTTATGCTCTCATGAGCGGATACATATTTGAATGATTTTGAATAAT 7141
Db 4954 AGCATTTTATCAGGGTTATGCTCTCATGAGCGGATACATATTTGAATGATTTTGAATAAT 5013
QY 7142 AAACAAATAGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACC 7201
Db 5014 AAACAAATAGGGTTCCGCGCACATTTTCCCGAAAAAGTGCCACCTGACGCTCTAAGAAACC 5073
QY 7202 ATTATTTATGATGATTAACCTTAAATAATAGGGTATACAGAGGCCCTTTCGCTCGCG 7261
Db 5074 ATTATTTATGATGATTAACCTTAAATAATAGGGTATACAGAGGCCCTTTCGCTCGCG 5133
QY 7262 CGTTTCGGTATGACGGTGAACCTCTGACACATGCGAGCTCCCGGAGAGCGGTCAAGCT 7321
Db 5134 CGTTTCGGTATGACGGTGAACCTCTGACACATGCGAGCTCCCGGAGAGCGGTCAAGCT 5193
QY 7322 TGCTGTAAAGCGGATCCCGGAGCAGCAAGACCCGCTCAGGGCGCTCAGGGGTGTGGC 7381
Db 5194 TGCTGTAAAGCGGATCCCGGAGCAGCAAGACCCGCTCAGGGCGCTCAGGGGTGTGGC 5253
QY 7382 GGGTGTCCGGGCTGGCTTAACTATGCGGCAATCAGAGCAGATTTGATCTGAGAGTCAACAT 7441

Db 5254 GGGTGTGGGGTGGCTTAACCTATGCGGCATCAGACGAGATTGCTAGAGATGCACCAT 5313

Qy 7442 ATGCGGTGTGAATACCGCACAGATCGTAAGGAGAAATACCGCATCAGG 7492

Db 5314 ATGCGGTGTGAATACCGCACAGATCGTAAGGAGAAATACCGCATCAGG 5364

RESULT 13

U00220

LOCUS

DEFINITION Human immunodeficiency virus type 1 Nef protein and neomycin phosphotransferase genes, complete cds.

ACCESSION U00220

VERSION

KEYWORDS U00220.1 GI:392790

SOURCE Human immunodeficiency virus 1 (HIV-1)

ORGANISM Human immunodeficiency virus 1

REFERENCE 1 (bases 1 to 6506)

AUTHORS Anderson, S., Shugars, D.C., Swannstrom, R. and Garcia, J.V.

TITLE Nef from primary isolates of human immunodeficiency virus type 1 suppresses surface CD4 expression in human and mouse T cells

JOURNAL J. Virol. 67 (8), 4923-4931 (1993)

MEDLINE 93323236

PUBMED 8331733

REFERENCE 2 (bases 1 to 6506)

AUTHORS Shugars, D.C.

TITLE Direct Submission

JOURNAL Submitted (11-AUG-1993) Diane C. Shugars, UNC-Lineberger Cancer Research Ctr, University of North Carolina at Chapel Hill, Manning Drive, Chapel Hill, NC 27599-7295 USA

FEATURES

source

1. .6506

/organism="Human immunodeficiency virus 1"

/proviral

/mol_type="unassigned RNA"

/db_xref="taxon:11676"

/clone="pLconsnefsN"

1855..2275

/codon_start=1

/products="HIV-1 consensus Nef protein"

/protein_id="AAA03700.1"

/db_xref="GI:392792"

/translation="MGGKWSKSVSGWPAVRMRRAEPAAGVGNVRDLKKGAIT
SNRTATNACWLAEQAEEVEGFPVRPQVPLRPMTYKAAVDLSHLKKEGLGLY
SKRQDILDWVYHTQGYFPDWNQYTPGGIRYPLTFGWCFLVPEPEKVEANE
NNCLLHPSQHGMDPEKEVLVWKFDKLAFLHMARELHPYYKDC"

2698..3492

/codon_start=1

/product="neomycin phosphotransferase"

/protein_id="AAA03699.1"

/db_xref="GI:392791"

/translation="MIEQDLHAGSPAAWVERLFGYDWAQQTIGCSDAVFRLSAQGR
PVLVFKTDLGSLNLEDAEALSLWLTGVPAAVLDVVTAGRDWLLGLBVPQDL
LSHLAPAEKVSIMADAMERLTLDPATCFPDQAKHLERARTMEAGLVDDDLDE
EQGLAPAEFLKARMPDGEDLVVTHGDACLPMIWNENGRFSFIDCRLGVADRY
QDIALATRIDABELGEWADRLVLYGIAAPSDSQRIAFYRLDLEFF"

ORIGIN

Query Match 40.6%; Score 3104.6; DB 14; Length 6506;

Best Local Similarity 70.8%; Pred. No. 0;

Matches 4793; Conservative 0; Mismatches 1474; Indels 504; Gaps 30;

Qy 549 TTTTGAAGACCCCACTAGTTTGGCAAGCTAGCTTAAGTAAGCCGCAATTTGCAAGG 608

Db 174 TTTTGAAGACCCCACTAGTTTGGCAAGCTAGCTTAAGTAAGCCGCAATTTGCAAGG 231

Qy 609 CATGG-AAAATACATACTGAGNATAGAGAGTTTCAGATCAAGTTAGGAACAGAGAC 667

Db 232 CATGGAAAATAATACATACTGAGNATAGAGAGTTTCAGATCAAGTTAGGAACAGAGAC 291

Qy 668 AGCAGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCC-GCTCAGGGCC 726

Db 292 AGCTGAATA---CCAAACAGGATATCTGTGTAAGCGGTCTCTGCCCCGGCTCAGGGCC 347

Qy 727 AGAACAGTTTGGAAACAGGAGAT-ATGGGCCAAACAGGATATCTGTGTAAGCAGTTCT 785

Db 348 AAGAACAGATGAGACAGCTGAGTGTATGGGCCAAACAGGATATCTGTGTAAGCAGTTCT 407

Qy 786 GCCCGGCTCAGGGCCAAAGAACAGATGGTCCACAGATGGTCCCGCTCAGCAGTTTC 845

Db 408 GCCCGGCTCAGGGCCAAAGAACAGATGGTCCACAGATGGTCCCGCTCAGCAGTTTC 467

Qy 846 TAGAAGAACCATCAGATGTTTCCAGGGTGCCTCCAAAGACCTG-AAATGACCTGTGCTT 904

Db 468 TAGTGAATCATCAGATGTTTCCAGGGTGCCTCCAAAGACCTGAAATGACCTGTACCTTA 527

Qy 905 TTTGAACATAACCATCAGTTTCGCTTCTGCTTCTGTTGCGCGCTTCTGCTCCCGAGCT 964

Db 528 TTTGAACATAACCATCAGTTTCGCTTCTGCTTCTGTTGCGCGCTTCCGCTTCCGAGCT 587

Qy 965 CAATAAAGAGCCCAACACCTCTACTCGGCGCGCAGTCTCCGATAGACTCGGTCGCC 1024

Db 588 CAATAAAGAGCCCAACACCTCTACTCGGCGCGCAGTCTCCGATAGACTCGGTCGCC 647

Qy 1025 CGGGTACCCGTATTCCTCAATAAAGCTCTTGTCTGTTGTCATCCGAATCGTGAACCTG 1084

Db 648 CGGGTACCCGTATTCCTCAATAAAGCTCTTGTCTGTTGTCATCCGAATCGTGTCTGCTG 707

Qy 1085 ATCTCTGGAGGCTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1144

Db 708 TTTCTTGGAGGCTCTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATG 767

Qy 1145 GGTTCACCGAGATTTGGAGACCTCTGCGCCAGGACCCAGCCGCCCGCGGAGGTA 1204

Db 768 GGTCTGTCGGATTTGGAGACCTCTGCGCCAGGACCCAGCCGCCCGCGGAGGTA 826

Qy 1205 AGCTGGCCAGCAACTTATCTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1264

Db 827 AGCTGGCCAGCAACTTATCTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 886

Qy 1265 CGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAAC 1324

Db 887 CGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGACCCGCTGGTGAAC 946

Qy 1325 GACGAGTTTCGGAACACCGCGCGCAACCTCTGGAGAGCTCCAGGAGACTTCGGGGCCGT 1384

Db 947 GACGAGTTTCGGAACACCGCGCGCAACCTCTGGAGAGCTCCAGGAGACTTCGGGGCCGT 1006

Qy 1385 TTTTGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTGGAGACTCTTTGGTGCACCCC 1444

Db 1007 TTTTGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTGGAGACTCTTTGGTGCACCCC 1057

Qy 1445 CTTAGAGAGGATATGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1504

Db 1058 -----GGATATGTTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTCCCGCTCC 1107

Qy 1505 GTCTGAATTTTGTCTTTCGTTTGGAGACCGGAGACCGCGCGCTGTTGT---CTGCTG 1561

Db 1108 GTCTGAATTTTGTCTTTCGTTTGGAAACGAGACCGCGCTCTTGTCTGCTGAGCGCTG 1167

Qy 1562 CAGCATCGTTCTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621

Db 1168 CAGCATCGTTCTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227

Qy 1622 CCGGGCTAGACTGTTTACACTCCCTTAAGTTTGAACCTTAGTCTACTGGAAGATGCTCA 1681

Db 1228 CC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGTCTACTGGAAGATGCTCA 1281

Qy 1682 GCGGATCGTCTCAACACCTGCTAGTCTCAAGAGAGACCTGTTGGTTACCTTCTGCTC 1741

Db 1282 GCGGATCGTCTCAACACCTGCTAGTCTCAAGAGAGACCTGTTGGTTACCTTCTGCTC 1341

Qy 1742 TGCAGAATGGCCCAACTTTTAACTCGGATGGCGCGGAGACGCACTTTTAAACCGAGACT 1801

Db 1342 TCGAGATGGCCAACTTTTAAGTCGGATGGCCGAGACGGCACCTTTAAACGAGACCT 1401
Qy 1802 CATCACCCAGGTTAAGATCAAGGCTTTTTCATCGGCCCGCATGGACACCCAGAGGT 1861
Db 1402 CATCACCCAGGTTAAGATCAAGGCTTTTTCATCGGCCCGCATGGACACCCAGAGGT 1461
Qy 1862 CCCCTACATCGTACCGGAGGCTTGGCTTTTGGACCCCGCTCGCTGGGTCAAGCCCTT 1921
Db 1462 CCCCTACATCGTACCGGAGGCTTGGCTTTTGGACCCCGCTCGCTGGGTCAAGCCCTT 1521
Qy 1922 TGTACACCTTAAGCTTCGGCTCTCTTCTCCATCGGCCCGCTCTCCCGCTTGAACC 1981
Db 1522 TGTACACCTTAAGCTTCGGCTCTCTTCTCCATCGGCCCGCTCTCCCGCTTGAACC 1581
Qy 1982 TCCTCGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTTAGGGC 2041
Db 1582 TCCTCGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTTAGGGC- 1640
Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGGCCCTTGTAAACTTCCCTGA 2101
Db 1641 -CCGAATTTCGCCACCATGGGTGGCAAGTGTCAAAACGTAGTGTAGTGGATGGCTGCT 1699
Qy 2102 CCCTGACATGACAAGAGTTACTTAACAGCCCTCTCTCTCCAACTCACTTTACAGGCTTCTAC 2161
Db 1700 GTAAGAGNAAGATGAGCGAGCTGAGCCACGACGAGGAGTGGAGCAG----- 1751
Qy 2162 TTAGTCCAGCACGAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAACAACTGGACCGA 2221
Db 1752 -----TATCTCGAGACCTCGAAACCATGGAGCAATTACAAGTAGCAATACAGACTA 1805
Qy 2222 CCGGTGTACTCTACCTTACCGAGTCCGGAGACAGTGTGGTCCGCGACACAGACT 2281
Db 1806 CCAATGTCTGTGTGCTCGCTGTAGAACCAAGAGGAGGAAGATGGGTTTTCCAGTCA 1865
Qy 2282 AAGAACCTTAGAACCTCGTGGAAAGGACCTTACACAGTCTCTGCTGACACCCCGCC 2341
Db 1866 GACCTCAGTAGCTTTAAGACCAATGACTTACAAGGAGAGTAGATCTTAGCCACTTTT 1925
Qy 2342 CTCAAAGTAGACGCATCG-----CAGCTTGATACAGCCCGCCCGCTGAAGGCTGCCGA 2397
Db 1926 TAAAGAAAGAGGGGGACTGGAAGGGTTAATTTACTCCCAAAAGACAAGATATCCTTG 1985
Qy 2398 CCCCGGGGTGGACATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTG 2457
Db 1986 ATCTGTGGGTCTACACACACAAGGCTACTTCCCTGATTTGGCAGAACTACACACAGGC 2045
Qy 2458 GTACCAACAGCTACAGTGTCCAATCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTG 2517
Db 2046 CAGGAATCAGATATCCACTGACCTTTTGGGTGCTCTCAAACTAGTACCAAGTTGAGCCAG 2105
Qy 2518 AGCCACGCGTGGTGACAGTGAACCATCACTGTAAAGGCCAGTCAAGGATGTGGGTACT 2577
Db 2106 AAGAGGTAGAAGAGGCCAATGAAGAGAGAACAACTGCTTGCTACACCTATGAGCCAGC 2165
Qy 2578 TCTGTAGTTGGTACCAAGAGAACCCAGGTAAAGCTCAAAGCTGTGATCTACTGACA 2637
Db 2166 ATGGGATGGATGACCCAGAGAAAGATTAATAGTGTGGAAGTTTGACAGCAAACTAGCAT 2225
Qy 2638 TCCACCCGGCACATGTTGTGCCAAGCAGATTACGCGGTAGCGGTACCGACTTC 2697
Db 2226 TTCAATCAGATGGCCCGAGAGCTGATCCGGAGTACTACAAGACTGCTGAGAAATTCGTTA 2285
Qy 2698 ACCTTCAACCATCAGCAGCCTCCAGCCAG-AGGACATCGCCACCTACTACTGCCAGCAATA 2756
Db 2286 ACTCAGGATCCGGCTGTGGNAATGTGTGATAGGTGTGGNAAGTCCCGAGGCTCCC 2345
Qy 2757 TAGCCTCTATGGTGTGTGGCCCAAGGGAACAAGTGGAAATCAAAAGAGTGGCTCAGG 2816
Db 2346 CAGCAGGAGCAAGATGCAAGAGCATGCATCTCAATTTAGTCAGCAACCAAGGTGTGGAAGT 2405
Qy 2817 ATCGGGTGGATCCGGCTCTGGTGGCTCAGGATCCGAGGTCCAATGTTGGAGAGCGGTGG 2876
Db 2406 CCCCAGGCTCCCCAGCGGAGCAAGATATGCAAAAGCATGCTCTCAATTTAGTCTCCAGAA 2465

Qy 2877 AGGTGTTGTCAACCTCGCCGGTCCCTCGCCTGTCTGCTCCGATCTCGCTTCGATTT 2936
Db 2466 TAGTCCCGCCCTAACTCCGCCCATCCCGCCCTAACTCCGCCAGTTCCGCCCATCTC 2525
Qy 2937 CACCACATATTTGGATGAGTTGGGTGAGACAGGACCTGGAAAAGGTCTTGAGTGGATTGG 2996
Db 2526 CGCCCATGCGTGAATAATTTTATTTATGAGAGGCGCGCCCTCGGCCTCTG 2585
Qy 2997 AGAAATTCATCCAGATAGCAGTACGATTAACATATGCGCCGTCTCTAAAGGATAGATTT- 3054
Db 2586 AGCTATTTCCAGAGTAGTGGAGGCTTTTGGAGGCTTAGGCTTTTGGCAAAAGCTTG 2645
Qy 3055 -ACAATATCCGAGACAACGCCAAGAACACATTTCTTCTGCAATGAGACACCTGAGACC 3113
Db 2646 GGCTGCAGGTCCGAGCGGATCTGATCAAGACAGAGATGAGATCGTTTCGATGATGA 2705
Qy 3114 CGAAGACACCGGGGTCTATTTTGTGCAAGCTTTACTTCGGCTTCCCTGGTTGCTTA 3173
Db 2706 ACAAGATGGATTGCACGACGTTCTCCGGCGCTTGGGTGGAGAGGCTATTCGGCTATGA 2765
Qy 3174 TTGGGGCAAGGGACCCCGGTACCGTCTCCAGTGTAAAGCCACACGACCGCCAGCGCC 3233
Db 2766 CTGGGCAACAGACAAATCGCTCTCTGATGCCCGCTGTTCCGGCTGTGAGCGAGGG 2825
Qy 3234 GCGACCAACCAACCGCGGCCCATCATCGGTTCGACGCCCTCTCTCGCCCGCAGAGGC 3293
Db 2826 GCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTCCCTGAAATGAATGCAGGACGA 2885
Qy 3294 GG-----CTCGGCGAGGGGGGGGGGAGTGACACAGAGGGGGCTGGACTT 3341
Db 2886 GCGACCGCGCTATCGTGGGTGGCCACGACGGGGTTCCTTGGCGAGTGTGCTCGACGT 2945
Qy 3342 CGCCTCGATCCAACTCTGCTACCTGCTGGATGGAATCCTCTTCTATGATGTGTCAT 3401
Db 2946 TGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGGGAAGTGCCTGGGCGAGGATCTCCT 3005
Qy 3402 TCTCACTGCTTGTCTCTGAGAGTGAAATTCAGCAGGAGCGCAGAGCC----- 3449
Db 3006 GTCATCTCACTTGTCTCTCGCGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCT 3065
Qy 3450 -----CCCGCGTACCMAGGAGGGCGAGAACAGCTCTATTAACGAGCTCAATCT 3497
Db 3066 GCATACGCTTGTATCCGGTACTGCCCATTCGACCAACAGCGAAACATCGCATCGAGC 3125
Qy 3498 AGACAGAGAGAGAGTACGATGTTTGGACAGAGACGTGGCGGGACCTTGAGATGG 3557
Db 3126 AGCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCA 3185
Qy 3558 GGGAAAGCCGAGAAAGAACCCCTCAGGAAGGCTGTACAATGAATGCAACTGCAGAAAGATAA 3617
Db 3186 GGG--GCTCGCGCCAGCCGAATGTTCCCGAGGCTCAAGGCGCATGCCCGCGGAG 3243
Qy 3618 GATGGCGGAGCCCTACGATGAGATTGGGATGAAAGGCGAGCGCGCGAGGGGCAAG--GG 3675
Db 3244 GATCTGCTGAGCCCATGGCGATGCTGCTGCGCAATATCATGTGGAAAATGGCGCG 3303
Qy 3676 CACGATGGCCCTTACAGGGTCTCAGTACACCAAGGACACCTAGAGCGCCCTTCAAC 3735
Db 3304 TTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGCGGACCGCTATCAGGACATAGCG 3363
Qy 3736 ATCAGCGCCTGCGCCCTCGCTAAC-----TCGACGCGCGCGCGATCCGGATTAGTCCAA 3791
Db 3364 TTGGCTACCGTGTATTTGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCGTG 3423
Qy 3792 TTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACAAATATCACAG 3851
Db 3424 CTTTACGTTATCGCGCTCCCGATTCGACGCGCATCGCTTCTATCGCTTCTTGAAGAG 3483
Qy 3852 CTGAAGCCTATAGATACGAGCCATAGATAAATAAAGATTTTATTAGTCTCCAGAA 3911
Db 3484 TTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAAGATTTTATTAGTCTCCAGAA 3543

QY 3912 AAGGGGGGAATGAAGACCCACCCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCATTT 3971
DB 3544 AAGGGGGGAATGAAGACCCACCCTGTAGTTTGGCAAGCTAGCTTAAGTAAACGCCATTT 3603
QY 3972 TGCAGGCGATGG-AAAATACATACTGAGAAATAGAGAACTTACAGATCAAGGTTAGGAACA 4030
DB 3604 TGCAGGCGATGGAAAAATACATACTGAGAAATAGAGAACTTACAGATCAAGGTTAGGAACA 3663
QY 4031 GA-GAGACAGCAGATATGGGCCCAACACAGGATATCTGTGTAAGCAAGTTCTGCCCCC-CC 4088
DB 3664 GATGGAACAGCTGAATATGGGCCCAACACAGGATATCTGTGTAAGCAAGTTCTGCCCCC 3723
QY 4089 TCAGGGCCCAAGAACAGTTCGGAACAGGAGAAATATGGGCCCAACACAGGATATCTGTGTAAGC 4148
DB 3724 TCAGGGCCCAAGAACAGTTCGGAACAGGAGAAATATGGGCCCAACACAGGATATCTGTGTAAGC 3783
QY 4149 AGTTCTGTCGGGCTCAGGGCCCAAGAACAGATGTTCCAGATGCGGTCGCCGCTCAG 4208
DB 3784 AGTTCTGTCGGGCTCAGGGCCCAAGAACAGATGTTCCAGATGCGGTCGCCGCTCAG 3843
QY 4209 CAGTTCTAGAGAACCATCAGATGTTCCAGGGTCCCAAGGACCTGGAATGACCCGT 4268
DB 3844 CAGTTCTAGAGAACCATCAGATGTTCCAGGGTCCCAAGGACCTGGAATGACCCGT 3903
QY 4269 GCCTATTTGAACCTAACCAATCAGTTGCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTCCC 4328
DB 3904 GCCTATTTGAACCTAACCAATCAGTTGCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTCCC 3963
QY 4329 CGAGCTCAATAAAGAGCCCAACACCCTCACTCGGCGCGCCAGTCTCCCGATAGACTGC 4388
DB 3964 CGAGCTCAATAAAGAGCCCAACACCCTCACTCGGCGCGCCAGTCTCCCGATAGACTGA 4023
QY 4389 GTCCGCGGGTACCGGTGTTCTCAATAAACCTCTTGCAGTTGTCATCCGACTCGTGGTCT 4448
DB 4024 GTCCGCGGGTACCGGTGTTCTCAATAAACCTCTTGCAGTTGTCATCCGACTCGTGGTCT 4083
QY 4449 CGCTGTTCTCGGAGGGTCT-CTCTGAGTGAATGACTACCCGTCAGCGGGTCTTTTCAG 4507
DB 4084 CGCTGTTCTCGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCGGGTCTTTTCAG 4143
QY 4508 TTTCTCCCACTACACAGCTCTCACTAACATTCCTGATGTCGCGCAGGACTCCGTCAGC 4567
DB 4144 TTTGGGGGCTGTCGCGGATCGGAGACCCCTGCCAGGAGCACCGACCCACCCAGGG 4203
QY 4568 CCGGTTTTGTTTATAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGACTACATCCT 4627
DB 4204 AGGTAAGCTG-----GCT 4216
QY 4628 GACTCTCGGCTTTATAAAGAAATGTTGAAGGGTCTGTGGACTATCTGCCACACGACTTT 4687
DB 4217 GCTCGGCGTTTCGGTGATGACGGTGAACAACTCTG-ACATGCGAGTCCCGGAGCG 4275
QY 4688 TTAAGATTTTATGCTCCTGGATGAGGATTTAGTCAATCTATCTCTGCTCTATTGTTGT 4747
DB 4276 GTCACAGCTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCGTCAGGCGCGTCAGCG 4335
QY 4748 GGGTTCTCGTATTTTAAATTTCTAGTTTGGACTCCCTTCTGAGAGCAGCGGATTGCA 4807
DB 4336 GGTGTTGGCG----- 4345
QY 4808 GAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGGTTGCTGGGCTCAGTGTGACA 4867
DB 4346 ----- 4345
QY 4868 TTTTGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGGAATCTAGAGTGCATA 4927
DB 4346 ----- 4345
QY 4928 CTCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGGATCTTTGCTCCTAGAGATCCTCT 4987
DB 4346 -----GGTGTGGGGCGCAGCCATGACCCAGTC----- 4373
QY 4988 ACGCCGACGATCGTGGCGGGGTACCGAGCTCGAATTCGTAATCATGTCATGCTGTT 5047

DB 4374 -----ACGTAGCCATAGCGGAGTGTATCTGGCTTAACATATGCGCATCAGAGCAGAT 4426
QY 5048 TCCTGTGTGAATTTGTTATCCGCTCACAAATCCACACATACAGCGCGGAGCATATAA 5107
DB 4427 TGTACTGAGA----- 4436
QY 5108 GTGTAAAGCCTGGGGTGCCTAATAGTGAGCTAACTCACATTAATTGGGTTGCGCTCACT 5167
DB 4437 ----- 4436
QY 5168 GCCCGCTTTCCAGTCGGGAAACCTGTGTCGAGCTGCAATTAATGAATCGGCCAACGCGC 5227
DB 4437 -----GTGCACCATATGCGGTGTGAAATACCCACAGATG 4471
QY 5228 GGGGAGAGCGGTTTGGCTATTTGGCGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGGC 5287
DB 4472 CGTAAGGAGAAATACCGCATCAGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGGC 4531
QY 5288 CTCGCTGCTCGGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATC 5347
DB 4532 CTCGCTGCTTCCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATC 4591
QY 5348 CACAGAATCAGGGGATAACGCGAGGAAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAG 5407
DB 4592 CACAGAATCAGGGGATAACGCGAGGAAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAG 4651
QY 5408 GAAACGTTAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCCGCCCTTACGAGCA 5467
DB 4652 GAAACGTTAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCCGCCCTTACGAGCA 4711
QY 5468 TCACAAAATCGACCTCAAGTCAGAGGTGGGAAACCCGACAGACTATATAAGATACCA 5527
DB 4712 TCACAAAATCGACCTCAAGTCAGAGGTGGGAAACCCGACAGACTATATAAGATACCA 4771
QY 5528 GCGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGG 5587
DB 4772 GCGCTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGCGCTTACCGG 4831
QY 5588 ATACCTGTCGGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAG 5647
DB 4832 ATACCTGTCGGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAG 4891
QY 5648 GTATCTCAGTTCCGTTAGGTGCTGTTCCAAAGTGGGCTGTGTGACGAACCCCGCT 5707
DB 4892 GTATCTCAGTTCCGTTAGGTGCTGTTCCAAAGTGGGCTGTGTGACGAACCCCGCT 4951
QY 5708 TCAGCCCGACCGCTCGCTTATCCGGTAACTATCTGTTGAGTCCAAACCCCGTAAAGCA 5767
DB 4952 TCAGCCCGACCGCTCGCTTATCCGGTAACTATCTGTTGAGTCCAAACCCCGTAAAGCA 5011
QY 5768 CGACTTATCGCCACTCGGACGACGACCTGGTAAACAGGATTAGCAGCGGAGGTATGTAGG 5827
DB 5012 CGACTTATCGCCACTCGGACGACGACCTGGTAAACAGGATTAGCAGCGGAGGTATGTAGG 5071
QY 5828 CCGTCTACAGAGTTCTTGAAGTGTGCTTAACTACGCTACACTAGAGGACAGTATT 5887
DB 5072 CCGTCTACAGAGTTCTTGAAGTGTGCTTAACTACGCTACACTAGAGGACAGTATT 5131
QY 5888 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC 5947
DB 5132 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC 5191
QY 5948 CGGCAAAACAAACCGCTGTAGCGGTGTTTTTTTGTTCAGCAGCAGCATACGCG 6007
DB 5192 CGGCAAAACAAACCGCTGTAGCGGTGTTTTTTTGTTCAGCAGCAGCATACGCG 5251
QY 6008 CAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG 6067
DB 5252 CAGAAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTG 5311
QY 6068 GAAACGAAACTCAGCTTAAGGATTTTGTTCATGAGATTATCAAAAGGATCTTCACTTA 6127

Db	5312	GAA	CGAAA	CTC	CAGT	TTAA	GGGAT	TTT	TGGT	CTAT	GAGAT	TAT	TCA	AAAA	AGGAT	CTT	CA	CGTA	5371		
Qy	6128	GAT	CTTT	TAA	ATTT	AAAA	AT	GAAG	TTTT	TAA	AT	CAAT	CTAA	AGTAT	TAT	TAT	GAGT	AAAA	CTTG	6187	
Db	5372	GAT	CTTT	TAA	ATTT	AAAA	AT	GAAG	TTTT	TAA	AT	CAAT	CTAA	AGTAT	TAT	TAT	GAGT	AAAA	CTTG	5431	
Qy	6188	GTC	GAC	AGT	TAC	CAAT	TGCT	TAA	TAC	GAG	GCA	CC	TAT	CTC	AGG	AT	CTGT	CTAT	TTCG	6247	
Db	5432	GTC	TG	CAG	AGT	TAC	CAAT	TGCT	TAA	TAC	GAG	GCA	CC	TAT	CTC	CAG	CGAT	CTGT	CTAT	TTCG	5491
Qy	6248	TT	CAT	CCAT	AGT	TGCT	GACT	CC	CGCT	CGT	CTAG	ATAC	TAC	TAC	GAT	ACG	GAG	GCGCT	TACC	6307	
Db	5492	TT	CAT	CCAT	AGT	TGCT	GACT	CC	CGCT	CGT	CTAG	ATAC	TAC	TAC	GAT	ACG	GAG	GCGCT	TACC	5551	
Qy	6308	AT	CT	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	6367	
Db	5552	AT	CT	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	5611	
Qy	6368	AG	CA	AT	AAA	CC	AG	C	AG	CG	CG	AG	CG	CG	AG	CG	CG	CG	CG	6427	
Db	5612	AG	CA	AT	AAA	CC	AG	C	AG	CG	CG	AG	CG	CG	AG	CG	CG	CG	CG	5671	
Qy	6428	CT	CA	T	CA	G	T	C	T	A	T	T	A	T	T	T	T	T	T	6487	
Db	5672	CT	CA	T	CA	G	T	C	T	A	T	T	A	T	T	T	T	T	T	5731	
Qy	6488	TT	T	G	C	G	CA	A	C	G	T	T	G	C	A	T	T	G	C	6546	
Db	5732	TT	T	G	C	G	CA	A	C	G	T	T	G	C	A	T	T	G	C	5791	
Qy	6547	G	C	T	T	C	A	T	T	C	A	T	T	C	C	C	A	A	T	6606	
Db	5792	G	C	T	T	C	A	T	T	C	A	T	T	C	C	C	A	A	T	5851	
Qy	6607	CA	AAAA	AG	CG	GT	AG	CT	CT	CG	CT	CG	AT	CG	GT	CT	CG	AG	6666		
Db	5852	CA	AAAA	AG	CG	GT	AG	CT	CT	CG	CT	CG	AT	CG	GT	CT	CG	AG	5911		
Qy	6667	G	T	T	A	T	C	A	C	T	C	A	T	T	C	T	C	T	6726		
Db	5912	G	T	T	A	T	C	A	C	T	C	A	T	T	C	T	C	T	5971		
Qy	6727	A	T	G	C	T	T	T	G	A	C	T	CA	CC	AA	G	T	CA	6786		
Db	5972	A	T	G	C	T	T	T	G	A	C	T	CA	CC	AA	G	T	CA	6031		
Qy	6787	A	C	C	G	A	G	T	G	C	T	CA	AT	AC	CG	CG	CA	CA	6846		
Db	6032	A	C	C	G	A	G	T	G	C	T	CA	AT	AC	CG	CG	CA	CA	6091		
Qy	6847	A	A	A	A	G	T	G	C	T	CA	AT	AC	CG	CG	CA	CA	CA	6906		
Db	6092	A	A	A	A	G	T	G	C	T	CA	AT	AC	CG	CG	CA	CA	CA	6151		
Qy	6907	G	T	T	G	A	T	C	G	A	T	CA	CC	AA	G	T	CA	CA	6966		
Db	6152	G	T	T	G	A	T	C	G	A	T	CA	CC	AA	G	T	CA	CA	6211		
Qy	6967	T	T	T	C	A	C	A	G	G	T	CA	AAAA	CA	G	G	A	A	7026		
Db	6212	T	T	T	C	A	C	A	G	G	T	CA	AAAA	CA	G	G	A	A	6271		
Qy	7027	A	A	G	G	G	C	A	C	G	A	A	T	G	T	CA	A	T	7086		
Db	6272	A	A	G	G	G	C														

QY	7207	TATCATGACATTAACTTATAAAATAAGGCGTATCAGAGGCCCTTTCGTCT	7257
DB	6452	TATCATGACATTAACTTATAAAATAAGGCGTATCAGAGGCCCTTTCGTCT	6502
RESULT 14			
BD138563	BD138563 6522 bp DNA linear PAT 18-SEP-2002		
LOCUS	Expression system of nucleotide with reduced immunogenicity for use		
DEFINITION	in gene therapy.		
ACCESSION	BD138563		
VERSION	BD138563.1 GI:232333508		
KEYWORDS	JP 2002508976-A/2.		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1 (bases 1 to 6522)		
AUTHORS	Radosevich, T.J. and Jr, C.J.L.		
TITLE	Expression system of nucleotide with reduced immunogenicity for use		
JOURNAL	in gene therapy		
COMMENT	Patent: JP 2002508976-A 2 26-MAR-2002; HUMAN GENE THERAPY RESEARCH INSTITUTE OS Artificial Sequence PN JP 2002508976-A/2 PD 26-MAR-2002 PF 13-JAN-1999 JP 2000540263 PR 14-JAN-1998 US 60/071409 PI THOMAS J RADOSEVICH, CHARLES J LINK JR PC C12N15/09, A61K31/711, A61K48/00, A61P37/06, C12N1/15, C12N1/19, PC C12N1/21. PC C12N5/10/A61K35/76, C12N15/09, C12R1/93, C12N15/00, C12N5/00, PC C12N15/00, C12R1/93) CC Description of Artificial Sequence : recombinant vector FH Key Location/Qualifiers FT source 1. .6522 FT /organism='Artificial Sequence'.		
FEATURES	Location/Qualifiers 1. .6522 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630"		
ORIGIN	Query Match 40.4%; Score 3096; DB 6; Length 6522; Best Local Similarity 70.8%; Pred. No. 0; Matches 4803; Conservative 0; Mismatches 1465; Indels 518; Gaps 31;		
QY	549	TTTTGAAAGACCCACCTGTAGCTGTTGGCAAGCTAGCTTAAGTAAGCCATTTTGCAAGG	608
DB	174	TTTTGAAAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAGTAAGCCATTTTGCAAGG	231
QY	609	CATGG-AAAATACATAACTGAGAAATAGAGAAGTTTCAGATCAAGTTAGGAACAGAGAC	667
DB	232	CATGGAAAAATACATAACTGAGAAATAGAGAAGTTTCAGATCAAGTCCAGGAACAAGAAAC	291
QY	668	AGCAGATATGGGCCCAACAGGATATCTGTGGTAAGCAGTTCCTGCGCCC-GCTCAGGGCC	726
DB	292	AGCTGGAATA---CCAAACAGGATATCTGTGGTAAGCGGTTCTGCCCCGGCTCAGGGCC	347
QY	727	AAGAACAGTTGGAAACAGGAGAAT-ATGGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCCT	785
DB	348	AAGAACAGATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCCT	407
QY	786	GCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGCGGTCCCGCCTTCAGCAGTTTC	845
DB	408	GCCCCGCTCAGGGCCAAAGACAGATGGTCCCCAGATGCGGTCCCGCCTTCAGCAGTTTC	467
QY	846	TAGAGACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTG-AAATGACCTGTGCTTA	904
DB	468	TAGTGAATCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAAAATGACCTGTGCTTA	527
QY	905	TTTGAACTAACCAATCAGTTTCGTCTTCGTGTTCTGTTCCGCGGCTTCGTGCTCCCGAGCT	964

Db	528	TTTTGAATCAACCAATCAGTTTCCTCTCGCTTCTGTTCTCGCGGCTTCCGCTCTCCGAGCT	587
Qy	965	CAATAAAGAGGCCCAACCCCTCACTCGCGCGCGCAGTCTCTCCGATAGACTGGTGGCC	1024
Db	588	CAATAAAGAGGCCCAACCCCTCACTCGCGCGCGCAGTCTCTCCGATAGACTGGTGGCC	647
Qy	1025	CGGGTACCCGGTATTTCCCAATAAAGCCCTCTGCTGTTGTCATCCGAATCGTGGACTCGTG	1084
Db	648	CGGGTACCCGGTATTTCCCAATAAAGCCCTCTGCTGTTGTCATCCGAATCGTGGCTCTCGTG	707
Qy	1085	ATCCTTTGGAGGGTCTCCTCAGATTGATGTCGCCACCTCGGGGGTCTTTCAATTGGGA	1144
Db	708	TTCTTTGGGAGGGTCTCCTCTCAGTGAATTGACTACCCACGACGGGGTCTTTTCATTGGG	767
Qy	1145	GGTTCCACCGAGATTGGAGACCCCTGCCAGGGACACCGACCCCGCGCGGGAGGTA	1204
Db	768	GGCTCTCGCGGATTTGGAGACCCCTGCCAGGGACACCGGA - CCCACACCGGGAGGTA	826
Qy	1205	AGCTGGCCAGCAACTTATCTGTGCTCTGCTCCGATTGCTAGTGTCTATGACTGATTTTATG	1264
Db	827	AGCTGGCCAGCAACTTATCTGTGCTCTGCTCCGATTGCTAGTGTCTATGTTGATGTTATG	886
Qy	1265	CGCTCGCTCGGTPACTAGTTAGCTAACTAGCTCTCTATCTGGCGGACCCGTGGTGGAACT	1324
Db	887	CGCTCGCTCGTACTAGTTAGCTAACTAGCTCTCTATCTGGCGGACCCGTGGTGGAACT	946
Qy	1325	GACGAGTTCCGGAACACCCCGCGCAACCTCGGAGACGTCGCCAGGACTTCGGGGCGGT	1384
Db	947	GACGAGTTCTGAACACCCCGCGCAACCCCTGGGAGACGTCGCCAGGACTTTGGGGCGGT	1006
Qy	1385	TTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTAGGACTCTTTGGTGCAACCCCC	1444
Db	1007	TTTTGTGGCCGACCTGAGGAGGGAGTGCATGTGGAATCCGACCCCGTCA - -----	1057
Qy	1445	CTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCCTCC	1504
Db	1058	-----GGATAATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCCTCC	1107
Qy	1505	GTCTGAATTTTGTCTTTCGGTTTGGGACCGAGCCGCGCGCTCTGT - --CTGCTG	1561
Db	1108	GTCTGAATTTTGTCTTTCGGTTTGGAAACCGAGACCGCGCGTCTGTGCTGCAAGCGCTG	1167
Qy	1562	CAGCATCTGTTGTTGTTCTGCTGACTGTTTCTGTTATTTGTTCTGAAATATGGG	1621
Db	1168	CAGCATCTGTTGTTGTTCTGTTCTGACTGTTTCTGTTATTTGTTCTGAAATATGGG	1227
Qy	1622	CCCGGCTAGACTGTTTACCCTCCCTTAAGTTTGAACCTTAGGTCACTCGGAAGATGCGA	1681
Db	1228	CC-----AGACTGTTTACCCTCCCTTAAGTTTGAACCTTAGGTCACTCGGAAGATGCGA	1281
Qy	1682	CGCGATCGCTCAACACGTCGGTAGATGTCAAGAGAGACGTTGGGTTTACTTCTGCTC	1741
Db	1282	CGCGATCGCTCAACACGTCGGTAGATGTCAAGAGAGACGTTGGGTTTACTTCTGCTC	1341
Qy	1742	TGCAGAAATGCCAAACCTTTAAAGTCGGAATGGCGCGGAGACGACCTTTTAAACCGAGACT	1801
Db	1342	TGCAGAAATGCCAAACCTTTAAAGTCGGAATGGCGCGGAGACGACCTTTTAAACCGAGACT	1401
Qy	1802	CATCACCAGGTTAAGATCAAGGTCCTTTTCACTCGCCCGCATGACACCCAGCAGGAGT	1861
Db	1402	CATCACCAGGTTAAGATCAAGGTCCTTTTCACTCGCCCGCATGACACCCAGCAGGAGT	1461
Qy	1862	CCCTATACGTGACCTGGGAGCCTTGGCTTTTGACCCCTCTCCCTGGGTCAAGCCCTT	1921
Db	1462	CCCTATACGTGACCTGGGAGCCTTGGCTTTTGACCCCTCTCCCTGGGTCAAGCCCTT	1521
Qy	1922	TGTACACCTTAAGCCTCCGCTCTCTTCCCTCATCCGCGCGCTCTCTCCCTTTGAACC	1981
Db	1522	TGTACACCTTAAGCCTCCGCTCTCTTCCCTCATCCGCGCGCTCTCTCCCTTTGAACC	1581
Qy	1982	TCCTCGTTGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCACTCTCTCTTAGGCGC	2041

Qy	3099	GGACAGCCTGAGACCCGAAACACACCCGGGTCTATTTTGTGCGAAGCCTTTACTTTCGGCTT	3158
Db	2707	GTTTTCGATGATTTGAAACAAGATGGATTCGACGCAAGGTTCCTCCGCGCGCTTGGGTGGAGAG	2766
Qy	3159	CCCCTGGTTTGTCTTATTTGGGCGCAAGGACCCCGGTCACTCGTCTCCAGTGTCTTAAGCCAC	3218
Db	2767	GCTATTTCGGCTATGACTTGGGCAACAACAGACAACTCGGTGTCTGTATGCCGCGTGTTCG	2826
Qy	3219	CACGACGCCAGCGCCGCGAACCAACAACCGGGGCCACCATCGCTTCGAGACCCCTGTGC	3278
Db	2827	GCTGTACGCGCAGGGGCGCCGGTTCCTTTTGTCAAGACCGACCTGTGCCGTGCCCTGAA	2886
Qy	3279	CCTGCGCCACAGAGCGG-----CTCGCCACGCGCGGGGGGCCAGTGCACAC	3326
Db	2887	TGAACTCGAGACGAGGACGCGCGGTATCGTGGCTGGCCACACGCGGCGTTCCTTTCGCG	2946
Qy	3327	GAGGGGCTGGACTTTCGCCCTGGATCCAAACTCTGTACTCTGTGATGGAATCTCTCTT	3386
Db	2947	AGCTGTGCTCGACGTTGTCACTGAAGCGGAGGGGACTGCGTGTCTATTGGCGAAGTGCC	3006
Qy	3387	CATCTATGGTGTCAATCTCACTGCGTTTGTCTGTAGAGTGAAGTTACGACGAGCGCAGA	3446
Db	3007	GGGCGAGGATCTCTGTCACTCACCTTGCTCTCTGCCGAGAAAGTATCCATCATGCGTGA	3066
Qy	3447	GCC-----CCCCGGCTACACAGCGGCCAGAACCAAGCTCTTA	3482
Db	3067	TGCAATGCGCGGTGTCATACGTTGTATCGGGTACTCGGCCATTCGACACCAACGAA	3126
Qy	3483	TACCGAGCTCAATCTAGGACGAGAGGAGTAGTACGATGTTTTCGACAAGAGAGCTGGCCG	3542
Db	3127	ACATCGCATCGACGAGCAGTACTCGGATGGAAGCCGCTCTTGTGATCAGAGATGATCT	3186
Qy	3543	GZACCTTGAGATGGGGGAAAGCCGAGAGAAACCCCTCAGAAAGCCGTGTACAATGA	3602
Db	3187	GGACGAGAGCATCAGGG--GCTCGCGCCACGCGCACTGTTCGCCAGGCTCAAGGCGCGC	3244
Qy	3603	ACTGCAGAAAGATAAGATGCGGGAGGCTACAGTGTAGATTTGGATGAAGCGCAGCGCG	3662
Db	3245	ATGCCGACGCGGAGGATCTGTCTGTGACCCATGGCGATGCCCTGTTCGCCAATATCATG	3304
Qy	3663	GAGGGGCAAG--GGCACGATGGCCTTTACGAGGCTCTCAGTACAGCCACCAAGGACACC	3720
Db	3305	GTGAAATGCGCGCTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC	3364
Qy	3721	TACGACGCCCTTCAATPGAGGCGCTGCCCCCTCGCTAAC-----TCGACGCGCGCGGGA	3776
Db	3365	TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGCGCAATGGGCT	3424
Qy	3777	TCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTGCCAGGCTCTAGTTTGGACT	3836
Db	3425	GACCGTTCTCTGTGCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTAT	3484
Qy	3837	CAACAAATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTA	3896
Db	3485	CGCCTTCTTGACAGTCTTCTGAGCGGACTCTGGGTTTCGATATAATAAAGATTTTA	3544
Qy	3897	TTTAGTCTCAGAAAAAGGGGGGAAATGAAGACCCCACTGTAGTTGGTTTGGCAAGCTAGCT	3956
Db	3545	TTTAGTCTCAGAAAAAGGGGGGAAATGAAGATCCCACTGTAGTTTGGCAAGCTAGCT	3604
Qy	3957	TAAGTAACGCATTTTGCAGGGCATG--AAAAATACATACTAGAGATAGAGAGTTTCAGA	4015
Db	3605	TAAGTAACGCCATTTTGCAGGGCATGAAAAATACATACTAGAGATAGAGAGTTTCAGA	3664
Qy	4016	TCAAGGTTAGGAAACAGA--GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGGTAAGC	4074
Db	3665	TCAAGGTCAGAAACAGATGAAACAGCTGATATGGGCCAAACAGGATATCTGTGGTAAGC	3724
Qy	4075	AGTTCTTGCCTCC--GCTCAGGGCCAGAAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGG	4133
Db	3725	AGTTCTTGCCTCCCGGCTCAGGGCCAGAAACAGATGGAAACAGCTGAATATGGGCCAAACAGG	3784

Qy	4134	ATATCTGTGGTAAGCAGTTCTCTGCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGATG	4193
Db	3785	ATATCTGTGGTAAGCAGTTCTCTGCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGATG	3844
Qy	4194	CGGTCCCGCCTCAGCAGTTTCTAGAGAAACATCAGATGTTTCCAGGGTGCCCCCAAGGAC	4253
Db	3845	CGGTCCCGCCTCAGCAGTTTCTAGAGAAACATCAGATGTTTCCAGGGTGCCCCCAAGGAC	3904
Qy	4254	CTGAAATGACCCCTGTGCTTATTTGAACTAACCAATCAGTTCGTTCTCGCTTCTGTTCG	4313
Db	3905	CTGAAATGACCCCTGTGCTTATTTGAACTAACCAATCAGTTCGTTCTCGCTTCTGTTCG	3964
Qy	4314	CGCGCTTCTGCTCCCGAGCTCAATAAAGAGCGCCACAACCCCTCACTCGCGCGGCCAGT	4373
Db	3965	CGCGCTTCTGCTCCCGAGCTCAATAAAGAGCGCCACAACCCCTCACTCGCGCGGCCAGT	4024
Qy	4374	CCTCCGATAGACTGCGTCCGCGGTACCGTGTTCTCAATAAACCCCTCTTCAGATTGCA	4433
Db	4025	CCTCCGATAGACTGCGTCCGCGGTACCGTGTTCTCAATAAACCCCTCTTCAGATTGCA	4084
Qy	4434	TCCGACTCGTGGTCTCGCTGTTCTTCTGGGAGGGTCT - CTCTGAGTGATTGACTACCGGTC	4492
Db	4085	TCCGACTCGTGGTCTCGCTGTTCTTCTGGGAGGGTCTCTCTGAGTGATTGACTACCGGTC	4144
Qy	4493	AGCGGGTCTTTTCAGTTTCTCCACCTCAACAGGTCTCAATAATTCCTGATGTGCGCGC	4552
Db	4145	AGCGGGGTTCTTTCAATTTGGGGGCTCGTCCGGATCGGAGACCCCTGCCAGGAGCCAC	4204
Qy	4553	AGGAGCTCCGTACGCCCGGTTTTTGTATTATAATAAATGCAAGAACAGTGTTCCTTCAA	4612
Db	4205	CGACCCACCAACCGGAGTTAAGCTG-----	4229
Qy	4613	GCCAGACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTAT	4672
Db	4230	-----GCTGCTCGCGGTTTCGGTGATGACGTGAAACCTCTG - ACACATG	4276
Qy	4673	CTGCCACACACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATC	4732
Db	4277	CAGTCCCGAGACGGTCAACAGCTGTGTCTGAAGCGGATGCCGGAGACACAAGCCCGT	4336
Qy	4733	CTCGTCTATTTTCTGCTCTCCGATTTTAAATTTCTAGTTTGCACTCCCTTCCTGAG	4792
Db	4337	CAGGCGCGTCAGCGGTGTGCGC-----	4361
Qy	4793	AGCACGCGATTCGACAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGTTGCGCT	4852
Db	4362	-----	4361
Qy	4853	GGGCTCAGTGTGAGATTTTGGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGG	4912
Db	4362	-----	4361
Qy	4913	AATCTAGATGCATCTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTT	4972
Db	4362	-----GGTGTCCGGGCGCAGCCATGACCCAGTC-----	4389
Qy	4973	GCCTAGAGATCCTCTACGCGGACGCATCGTGCCGGGTACCGAGCTCGAAATTCGTAATC	5032
Db	4390	-----ACGTAGCGATAGCGGAGTGATATCTGGCTTAACTATGC	4427
Qy	5033	ATGGTCATAGCTGTTCTGTGTGAATTTGTTATCGCTCAACAATTCACAACAATACG	5092
Db	4428	GGCATCAGACGAGATTGTACTGAGA-----	4452
Qy	5093	AGCGGAGCATAAAGTGTAAACCTCGGGTGCTTAATGATGAGTACTACATTAAT	5152
Db	4453	-----	4452
Qy	5153	TGCGTTGCGCTCACTGCCCGCTTTCCAGTCCGGAACCTGTCTGTGCCAGCTGCATTAATG	5212
Db	4453	-----GTGCACCATATGCGGTGTGA	4472
Qy	5213	AATCGGCCAAACGCGCGGGAGAGCGGTTTTCGATATTGGGCGCTCTTCGCGTCTTCGCT	5272

Db 4473 AATACCGCAGATGCGTAAGAGAAATACCGCATCAGCGCTCTTCCGCTTCTCGCT 4532
Qy 5273 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5332
Db 4533 CACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4592
Qy 5333 GGTAAATACGGTTATCCACAGAAATCAGGGATTAACCGAGAAAGAACATGTGAGCAAAAGG 5392
Db 4593 GGTAAATACGGTTATCCACAGAAATCAGGGATTAACCGAGAAAGAACATGTGAGCAAAAGG 4652
Qy 5393 CCAGCAAAAGGCGAGAAACCGTAAAGGCGCGTGTGCTGGGCTTTTCCATAGGCTCGG 5452
Db 4653 CCAGCAAAAGGCGAGAAACCGTAAAGGCGCGTGTGCTGGGCTTTTCCATAGGCTCGG 4712
Qy 5453 CCCCTGTAGCAGATCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAAACCGGACAGG 5512
Db 4713 CCCCTGTAGCAGATCATCAAAAATCGAGCTCAAGTCAAGAGTGGCGAAACCGGACAGG 4772
Qy 5513 ACTATAAGATACCGAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGAC 5572
Db 4773 ACTATAAGATACCGAGCGTGTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTCCGAC 4832
Qy 5573 CTTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTTCTCA 5632
Db 4833 CTTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGAAGCGTGGCGCTTTCTCA 4892
Qy 5633 TAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTTCGTTCCGTCAGCTGGCTGTGT 5692
Db 4893 TAGCTCACGCTGTAGTATCTCAGTTCGGTGTAGTTCGTTCCGTCAGCTGGCTGTGT 4952
Qy 5693 GCACGAACCCCGTTACCGCGACCGCTGCGCTTATCCGCTTAATCTGCTTGTAGTC 5752
Db 4953 GCACGAACCCCGTTACCGCGACCGCTGCGCTTATCCGCTTAATCTGCTTGTAGTC 5012
Qy 5753 CAACCCGGTAAGACAGCTTATCGCACTGGCAGCAGCACTGGTAAACAGGATAGCAG 5812
Db 5013 CAACCCGGTAAGACAGCTTATCGCACTGGCAGCAGCACTGGTAAACAGGATAGCAG 5072
Qy 5813 AGCGAGTATGTAAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAATCTAGCGCTAC 5872
Db 5073 AGCGAGTATGTAAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAATCTAGCGCTAC 5132
Qy 5873 TAGAAGCAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTCGGAAAGAGCT 5932
Db 5133 TAGAAGCAGTATTTGGTATCTGCGCTCTGCTGAGCCAGTTACCTCGGAAAGAGCT 5192
Qy 5933 TGGTAGCTCTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGGTTTTTTTGTGCAA 5992
Db 5193 TGGTAGCTCTTGATCCGGCAAAACAAACCGCTGGTAGCGGTGGTTTTTTTGTGCAA 5252
Qy 5993 GCAGCAGATACCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGG 6052
Db 5253 GCAGCAGATACCGCAGAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGGG 5312
Qy 6053 GTCTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGAGATTAACA 6112
Db 5313 GTCTGACGCTCAGTGGAACGAAACCTCAGTTAAGGATTTTGGTCAATGAGATTAACA 5372
Qy 6113 AAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 6172
Db 5373 AAGGATCTTACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 5432
Qy 6173 ATATAGTAAATCTTGTGTGACAGTTACCAATGCTTAATCAGTAGGCACTTATCTCAGC 6232
Db 5433 ATATAGTAAATCTTGTGTGACAGTTACCAATGCTTAATCAGTAGGCACTTATCTCAGC 5492
Qy 6233 GATCTGCTATTTCTGTTTCATCATAGTTCGCTGACTCCCGCTGCTGATCACTACGAT 6292
Db 5493 GATCTGCTATTTCTGTTTCATCATAGTTCGCTGACTCCCGCTGCTGATCACTACGAT 5552
Qy 6293 ACGGAGGGCTTACCAATCTGGCCCCAGTGTGCAATGTATACCGCGAGACCCAGCTCACC 6352

Db 5553 ACGGAGGGCTTACCAATCTGGCCCCAGTGTGCAATGTATACCGCGAGACCCAGCTCACC 5612
Qy 6353 GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCCGAGAAGTGGTCC 6412
Db 5613 GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCCGAGAAGTGGTCC 5672
Qy 6413 TGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAGCTTAGAGTAAGTAG 6472
Db 5673 TGCAACTTTATCCGCTCCATCCAGTCTATTAAATTTGTTCCCGGAAGCTTAGAGTAAGTAG 5732
Qy 6473 TTGCGCAGTTAATAGTTTGGCAACGTTTGGCCATGCTACAGGC - TGTGTTGTTGACG 6531
Db 5733 TTGCGCAGTTAATAGTTTGGCAACGTTTGGCCATGCTACAGGCATCGTGGTGTGACG 5792
Qy 6532 CTCGTCGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATG 6591
Db 5793 CTCGTCGTTTGGTATGGCTTCAATTCAGCTCCGTTCCCAACGATCAAGCGAGTTACATG 5852
Qy 6592 ATCCCCCATTTGTTGCAAAAAGCGTTAGCTCTTTCGGTCTCCGATCGTTGTGTCAGAAG 6651
Db 5853 ATCCCCCATTTGTTGCAAAAAGCGTTAGCTCTTTCGGTCTCCGATCGTTGTGTCAGAAG 5912
Qy 6652 TAAGTTGGCCGAGTGTATCACTCATGTTTATGCGAGCAGTGCATTAATCTTCTACTGT 6711
Db 5913 TAAGTTGGCCGAGTGTATCACTCATGTTTATGCGAGCAGTGCATTAATCTTCTACTGT 5972
Qy 6712 CATGCCATCCGTAAGATGCTTTCTGACGTGTTGAGTACTCAACCAAGTCAATCTGAGA 6771
Db 5973 CATGCCATCCGTAAGATGCTTTCTGACGTGTTGAGTACTCAACCAAGTCAATCTGAGA 6032
Qy 6772 ATAGTGTATGCGGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGATTAATACCGCGCC 6831
Db 6033 ATAGTGTATGCGGCGACCGAGTTGCTCTTCCCGCGCTCAACACGGGATTAATACCGCGCC 6092
Qy 6832 ACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAGAGTCTTTCGGGCGGAAACTCTC 6891
Db 6093 ACATAGCAGAACTTTTAAAGTGTCTCATCATTTGGAAGAGTCTTTCGGGCGGAAACTCTC 6152
Qy 6892 AAGGATCTTACCGCTTGGAGTCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATC 6951
Db 6153 AAGGATCTTACCGCTTGGAGTCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATC 6212
Qy 6952 TTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC 7011
Db 6213 TTACGATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC 6272
Qy 7012 CGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATATCTCATACTTCTCTTTTCA 7071
Db 6273 CGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATATCTCATACTTCTCTTTTCA 6332
Qy 7072 ATATTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGAT 7131
Db 6333 ATATTATTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGAT 6392
Qy 7132 TTAGAAAATAAACAATAAGGGTTCCCGCAGCATTTCCCCGAAAGTGCCACCTGACGT 7191
Db 6393 TTAGAAAATAAACAATAAGGGTTCCCGCAGCATTTCCCCGAAAGTGCCACCTGACGT 6452
Qy 7192 CTAGAAACCATTTATTCATGACATTAACCTTATAAAAATAGCGGTATCACGAGGCCCTT 7251
Db 6453 CTAGAAACCATTTATTCATGACATTAACCTTATAAAAATAGCGGTATCACGAGGCCCTT 6512
Qy 7252 TCGTCT 7257
Db 6513 TCGTCT 6518

RESULT 15
SYNMOVI
LOCUS
DEFINITION
ACCESSION

SYNMOVI 6374 bp DNA linear SYN 27-APR-1993
Moloney murine leukemia virus retroviral vector pUXSHD, complete
sequence.
M64753


```

VERSION      M64753.1 GI:208855
KEYWORDS     cloning vector; hisicidinol dehydrogenase; retroviral vector.
SOURCE       unidentified cloning vector
ORGANISM     other sequences; artificial sequences; vectors.
REFERENCE    1 (bases 1 to 6374)
AUTHORS      Stockschlaeder, M.A., Storb, R., Osborne, W.R. and Miller, A.D.
TITLE        L-hisicidinol provides effective selection of
              retrovirus-vector-transduced keratinocytes without impairing their
              proliferative potential
              Hum. Gene Ther. 2 (1), 33-39 (1991)
JOURNAL      91322161
MEDLINE      1650586
PUBMED
COMMENT      Original source text: Synthetic, DNA.
FEATURES     location/Qualifiers
              1..6374
               /organism="unidentified cloning vector"
               /mol_type="genomic DNA"
               /db_xref="taxon:45196"
              175..763
               /note="Mo-MuSV 5' long terminal repeat"
              833..1642
               /note="extended packaging signal"
              1223..1225
               /note="gag ATG start codon to TAG stop codon"
              1655..1994
               /note="simian virus 40 early promoter"
              2004..3308
               /EC_number="1.1.1.23"
               /codon_start=1
               /transl_table=11
               /product="histidinol dehydrogenase"
               /protein_id="AAA73023.1"
               /db_xref="GI:208856"
               /translation="MSFNLLIDWNSCSPEQALLTRPAISASISITRTVSDILDNVK
              TRGDALREYSAKEDTEVTRALVTPPEIAAAGARLDELQMTAAVKNIEPHSAQ
              TLPVDVETQGVRCQVTRVSVGLYIPGSSAPLFLVLMATPARIAGCQKVLIC
              RPPADIELLYAALQGVQIEIFNVGGQAIAAALAFGSEVPKVDKIPFPGNAFVTEAK
              SQVSRDLGAAIDMPAGSEVLVIADSGATPDFVASDLLSOAHEGPDVQVILLTPDAD
              IARKVAVEROLAEPRADTAROLASRLIVTKDLAOCVAISNOYGPBHLIIOTRN
              ARDLVDALISNGSVPLGDWSPESAGDYASGNHVLPTVGYTATCSSLGLADPQKRMV
              QELSKAGFSALASITETLAAERUTAHNAVTLRVNALKEQA"
              3420..4013
               /note="Mo-MuLV 3' long terminal repeat"

ORIGIN
Query Match      40.4%; Score 3094.8; DB 12; Length 6374;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 4802; Conservative 0; Mismatches 1382; Indels 538; Gaps 34;

Qy 549 TTTTGAAGAGCCACACCTGAGTTTGGCAAGCTAGCTTAAAGTAACGCCATTTTGCAGG 608
Db 174 TTTTGAAGAGCCACACCGTAGG--TGGCAAGCTAGCTTAAAGTAACGCCATTTTGCAGG 231

Qy 609 CATGG-AAAATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGACACAGAGAC 667
Db 232 CATGAAAAATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGACACAGAGAAC 291

Qy 668 AGCAGATATGGGCCAAACAGAGATATCTGTGTAGAGATTCCTGCCCC--GCTCAGGGCC 726
Db 292 AGCTGAATA-----CCAAACAGGATATCTGTGTAGAGATTCCTGCCCCGCTCAGGGCC 347

Qy 727 AAGAACAGTTGGAACAGGAGAT-ATGGGCCAAACAGGATATCTGTGTAGAGATTCCT 785
Db 348 AAGAACAGATGAGACAGCTGAGTATGGGCCAAACAGGATATCTGTGTAGAGATTCCT 407

Qy 786 GCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGTTTC 845
Db 408 GCCCGGCTCAGGGCCAAAGACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGTTTC 467

Qy 846 TAGAAGACCATCAGATGTTTCCAGGTGCCCCAAGGACCTG-AAATGACCTGTGCGCTTA 904
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTGGCCCCAAGGACCTGAAAAATGACCTGTACCTTA 527

```

```

Qy 905 TTTGAACATAACCAATCAGTTTGGCTTCTCGCTTCTGTTTCGGCGCGCTTCTGCTCCCGAGCT 964
Db 528 TTTGAACATAACCAATCAGTTTGGCTTCTCGCTTCTGTTTCGGCGCGCTTCTGCTCCCGAGCT 587
Qy 965 CAATAAAAGAGAGCCACAAACCCCTCAGTCGGCGGCCAGTCTCCGATAGAGTGGGTGCGCC 1024
Db 588 CAATAAAAGAGAGCCACAAACCCCTCAGTCGGCGGCCAGTCTTCCGATAGAGTGGGTGCGCC 647
Qy 1025 CGGTACCCGCTATTCCCAATAAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGACTCGCTG 1084
Db 648 CGGTACCCGCTATTCCCAATAAAAGCCCTTCTGCTGTTTGCATCCGAATCGTGCTCGCTG 707
Qy 1085 ATCTTTGGAGGGTCTCTCTCAGATTGATTGATGCCACCTCGGGGGTCTTTTCATTGGA 1144
Db 708 TTCTTTGGAGGGTCTCTCTCAGATTGATTGATGCCACCTCGGGGGTCTTTTCATTGGA 767
Qy 1145 GGTTCACACCGAGATTTCGAGACCCCTGCCAGGAGACACCGACCCCGCCCGCGGAGGTA 1204
Db 768 GGCTCGTCCGGGATTTGGAGACCCCTGCCAGGAGACACCGA-CCACACCGGAGGTA 826
Qy 1205 AGCTGGCCAGCACTTATCTGTCTGTCGCAATTGTCTAGTGTCTATGACTGATTTTATG 1264
Db 827 AGCTGGCCAGCACTTATCTGTCTGTCGCAATTGTCTAGTGTCTATGTTGATGTTATG 886
Qy 1265 CGCTCGCTCGGTACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGGAACT 1324
Db 887 CGCTCGCTCTGACTAGTTAGTAACTAGCTCTGTATCTGGCGGACCCGCTGGTGGAACT 946
Qy 1325 GACGAGTTCCGAACACCCGCGCGCAACCTGGGAGAGTCCACAGGACTTCGGGGGCGGT 1384
Db 947 GACGAGTTCTGAACACCCGCGCGCAACCTGGGAGAGTCCACAGGACTTCGGGGGCGGT 1006
Qy 1385 TTTTGTGGCCGCACTCAGTCTTAAATCCGATCGTTTGGAGTCTTTGGTGCAACCCCC 1444
Db 1007 TTTTGTGGCCGCACTCAGGAGAGGAGTCTGATGGAATCGGACCC----- 1052
Qy 1445 CTTAGAGAGGAGATATGTGTTCTGTGTAGAGACGAGAACCTAAACAGTTTCCCGCTCC 1504
Db 1053 -----CGTCAGGATATGTGTTCTGTGTAGAGACGAGAACCTAAACAGTTTCCCGCTCC 1107
Qy 1505 GTCTGAATTTTGTGTTTGGTTGGGACCGAGCCGCGCGCGCTTGT---CTGCTG 1561
Db 1108 GTCTGAATTTTGTGTTTGGTTTGGAAACCGAGCCGCGCGCTTGTCTGCTGCAAGCGCTG 1167
Qy 1562 CAGCATGTTCTGTGTTGTCTCTGCTGACTGTCTTCTGTTTCTGTTTCTGTAATAATATCGG 1621
Db 1168 CAGCATGTTCTGTGTTGTCTCTGCTGACTGTCTTCTGTTTCTGTTTCTGTAATAATATCGG 1227
Qy 1622 CCCGGGCTAGACTGTTTACCCTCCCTTTAAGTTTGTGACCTTAGGTCACTGGAAAGATGTCGA 1681
Db 1228 CC-----AGACTGTTACCCTCCCTTTAAGTTTGTGACCTTAGGTCACTGGAAAGATGTCGA 1281
Qy 1682 GCGATCGCTCAAAACAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACCTTCTGCTC 1741
Db 1282 GCGATCGCTCAAAACAGTCGGTAGATGTCAAGAGAGAGCTTGGGTACCTTCTGCTC 1341
Qy 1742 TGCAGATGCGCAACCTTTTAAACGTGCGATGCGCCGAGAGCGGACCTTTTAAACCGAGACT 1801
Db 1342 TGCAGATGCGCAACCTTTTAAACGTGCGATGCGCCGAGAGCGGACCTTTTAAACCGAGACT 1401
Qy 1802 CATCACCCAGGTTAAGATCAAGGTCTTTTCACTTGGCCCGCATGGAGACCCAGACCCAGGT 1861
Db 1402 CATCACCCAGGTTAAGATCAAGGTCTTTTCACTTGGCCCGCATGGAGACCCAGACCCAGGT 1461
Qy 1862 CCCCTACATCGTACCTGGGAAGCTTGGCTTTTGAACCCCTCTCCCTGGGTCAAGCCCTT 1921
Db 1462 CCCCTACATCGTACCTGGGAAGCTTGGCTTTTGAACCCCTCTCCCTGGGTCAAGCCCTT 1521
Qy 1922 TGTAACCCCTAAGCTCGGCTCTCTTCTCTCAATCGCCCGCTCTCTCCCTTGAACC 1981
Db 1522 TGTAACCCCTAAGCTCGGCTCTCTTCTCTCAATCGCCCGCTCTCTCTCCCTTGAACC 1581

```

1982 TCCTCGTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCACTCTCTCTAGCGC 2041
Db |||||
1582 TCCTCGTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCACTCTCTCTAGCGC 1641
Qy |||||
2042 CCCCATATGCCCATATGAGATCTTATATGGGGACCCCGCCCTTGTAAATCTCCCTGA 2101
Db |||||
1642 CGGAATTCGTAACTCGAGGATCCGGCTGTGGAATGTGTCTCAGTTAGGTGTGGAAGT 1701
Qy |||||
2102 CCTGTACATGACAAGAGTTACTTAACAGCCCTCTCTCCAAAGCTCACTTACAGGCTTCTAC 2161
Db |||||
1702 CCCAGGCTCCCGAG-----CAGGCAGAAAGTATGAAAGATGCAATCTCA 1746
Qy |||||
2162 TTAGTCCAGCAGGAAGTCTGGAGACCTCTGGGGCAGCCTTACCAAGAACTGACCGCA 2221
Db |||||
1747 ATTAGTCAGAACCCAGGTGTGGAAGTCCCGAGGCTCCCGAGGCGCAGAGTATGCARA 1806
Qy |||||
2222 CCGGTGGTACTCACCTTTACCGAGTGGCGGACACAGTGTGGTTCGCCGACCAAGACT 2281
Db |||||
1807 GCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCC----- 1861
Qy |||||
2282 AGAACCTAGAACCTCGCTGGAAGACCTTACACAGTCTCTGCTGACACCCCGCCCGC 2341
Db |||||
1862 --GCCCTTAACCTCCGCCAGTTCCGCCCAATCTCCGCCCACTGCTGACATAATTTTTTT 1919
Qy |||||
2342 CTCAAAGTAGAGCGCATCGAGCTTGGATACACGCGCCGACAGTGAAGGCTGCCGACCCC 2401
Db |||||
1920 ATTTATGACAGGCCGAGCGCGCT-----CGGCTCTGAGCTATTCAGAGTA 1969
Qy |||||
2402 GGGGTGGACCATCTCTTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTGGTAG 2461
Db |||||
1970 GTGAGGAGGCTTTTGGAGGAGATCCCGGACCATGAGCTCAATACCTGANTGACTG 2029
Qy |||||
2462 CAACAGCTACAGGTCTCACTCCGACATCCAGTACACAGCCAGAGCCCAAGCAGCTGAGG 2521
Db |||||
2030 GAACAGCTAGCGCTGAA-----CAGCAGGTGCGGTGCTGACGCTCCGGGA 2079
Qy |||||
2522 CCAGCTGGGTGACAGAGTGACCATCACCTGTGAAGCCAGTCAGGATGGGTACTTCTG 2581
Db |||||
2080 TTTCCGCTCTGACAGTATTACCGGACGGTCAGCGATATTCTGGATATGTAAACGC 2139
Qy |||||
2582 TAGCTTGTGACAGAGAGCCAGGTAAAGCTTCAAGCTGTGTATCTGACGACATCA 2641
Db |||||
2140 GCGGTGACGATGCTGCTGCTGAATACAGCGCTAAATTTGATAAAACAGAAAGTGACAGCG 2199
Qy |||||
2642 CCGGCACACTGGTGTGCCAACAGATTCAGCGGTAGCGGTAGCGGTACGACTTCACT 2701
Db |||||
2200 TAGCGTCACTCCCTGAAGAGATCGCGCGCGCGCGCGCTGAGACGACGAATTA----- 2256
Qy |||||
2702 TCACCATCAGCAGCTCCAGCAGAGGACATCGCACCTACTTCTGCGCAATATAGCC 2761
Db |||||
2257 --AACAGGCGATGACCGCTGCGCTCMAAATATTGAAACGTTCCATTCCGCGAGACGCT 2314
Qy |||||
2762 TCTATCGGTGTTGGGCCAAGGACCAAGGTG- GAAATCAAAACGAGGTGGCTCAGGATCG 2820
Db |||||
2315 ACCGCTGTAGATGTGGAACCCAGCAGCGGTGCGTTGCCAGCAGTTTACGCGTCCCGT 2374
Qy |||||
2821 GGTGATCCGGTCTGCTGGCTCAGGATCGAGGTCCAACTGTTGGAGAGCGGTGAGGT 2880
Db |||||
2375 CTCGCTGTGCTGTATATTCCCGGGGCTCGGCTCCGCTCTTCTCAACGGTGTCTGAT 2434
Qy |||||
2881 GTTGTCAACCTTGGCGGTCCCTGCGCGTGTCTGCTCGCATCTGGCTTCGATTTCACT 2940
Db |||||
2435 GCTGGCGACCGCGCGCATTTGGGGATGCCAGAGGTGTTCTGTCTCCCGCGCC 2494
Qy |||||
2941 ACATATTTGATGAGTTGGGTGAGACAGGACCTTGGAAAAGGCTTTGAGTGGATTGGAGAA 3000
Db |||||
2495 CATCGCTGATGAATCTCTATGCGCGCAAC-----TGTTGGGTGCGAGGA 2542
Qy |||||
3001 ATTATCCAGATAGCAGTACGATTAATCTATGCGCGTCTCTAAAGATAGATTACAATA 3060
Db |||||
2543 AATCTTTAAGTGGCGCGCGAGCGGATTCGCGCTCT-----GGCC 2585
Qy |||||
3061 TCGCGAGACAAAGCCAAAGACATTTGTTCTTGCAATGGACAGCCTGAGACCCGAGAC 3120

2586 TTCGCGAGCGAGTCCGTACCGAAAGTGATAAATTTTGGCCCGGCAACGCTTTGTA 2645
Qy |||||
3121 ACCGGGTCTATTTTGTGCAAGCTTACTTTCGGCTTCCCTGGTTGCTTATTTGGGGC 3180
Db |||||
2646 ACCGAAGCCAAACGTCAGGTCCAGCGAGCGCTCTCGACGGCGGGCTATCGAT----- 2696
Qy |||||
3181 CAAGGGACCCCGGTCACTCTCCAGTCTTAAGCCACACAGCAGCCAGCGCGGACCA 3240
Db |||||
2697 -ATGCCAGCGCGCGCTCTGAAGTACTGATGATCGCA-GACAGCGCGCAACACCGGAT 2754
Qy |||||
3241 CCAACACCGCGGCCCAACCATCGCTCGCAGCCCTGTCTCTCGGCCCAAGAGCGGCTCG 3300
Db |||||
2755 TGTGCTGCTTCTGACCTGCTCTCCAGGCTGAGCACGGCCGATTTCCAGGTGATCTCTC 2814
Qy |||||
3301 CCAGCGCGGGGGGGCGAGTGACACAGAGGGGGCTGGACTTCGCCCTGGATCCCAACTC 3360
Db |||||
2815 TGACGCTGATGCTGACATTTGCCGAAGTGGCGG-----AGGCGGTAGAACGTCACATG 2870
Qy |||||
3361 TGCTACCTGCTGGATGGAATCTCTTCTATGTGTCTATCTCACTGCTTGTTCCTG 3420
Db |||||
2871 GCGGAATGCTCCGCGCGGACACCCCGCGGAGGCCCTGAGCGCAGTCTGATGTG 2930
Qy |||||
3421 AGAGTGAAGTTTCCAGCAGGAGCGCAGACCCCGCGGTACTACAGCAGCGCCAGAACAGCTC 3480
Db |||||
2931 ACCAAAGATTTAGCGCAGTGGCTG- --CCATCTCTAATCAGTATGGCGCGCACTTA 2987
Qy |||||
3481 TATAAGAGCTCAATCTAGGACGAGAGAGTACGATGTTTGGACAGAGAGCTGGC 3540
Db |||||
2988 ATCATCCAGACGCGCAATCGCGCGATTTGGTGGATGCGATTTACAGCGCAGGCTCGTA 3047
Qy |||||
3541 CGGACCTGCTGAGTGGGGGAAAGCCGAGAGGAAGAACCTTCAGGAAGGCTGTACAT 3600
Db |||||
3048 TTTCTCGGACCTGCTGCGCGAATCCGCGGTGATTTAGCTTCCGGAACCAACCATGTT 3107
Qy |||||
3601 GAACTGCAGAAAGATGAAGTGGCGAGGCTCACTAGTGAGATTTGGATGAAAGCGAGCGC 3660
Db |||||
3108 TTACCGACTATGGCTATCTGCTACCTGTTCCAGCCTTGGGTAGCGGATTTCCAGAAA 3167
Qy |||||
3661 CGGAGGGCAGGGGCGAGTGGCTTTACAGGGTCTCAGTACAGCCACAGAGACACC 3720
Db |||||
3168 CGGATGACCGTTTCAGGA- ACTGTGAAAAGCGGGCTTTTCCGCTCTGGCATCAACATTGA 3226
Qy |||||
3721 TACGACGCTTTCACATGACGCGCTGCGCCCTGCTAACTCGACGCGCGCGGATCCG 3780
Db |||||
3227 AACATTGGCGCGCGAGAACGCTGACCGCCCTAATAAATGCGGTGACCTCGCGGTAAA 3286
Qy |||||
3781 GATTAGTCCAATTTGTTAAAGACGAGATATCAGTGGTCTCAGGCTCTAGTTTGACTCAAC 3840
Db |||||
3287 CGCCCTCAAGAGCAAGCATGAGCACTGAAACACTCTCAGCGTCTGCTGACTTAGCCGG 3346
Qy |||||
3841 AATATCACAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900
Db |||||
3347 GATCTGCC-----TAGGCTTTTGCAAAAAGCTTATCGATAAATAAAGATTTTATTTA 3400
Qy |||||
3901 GTCTCCAGAAAAGGGGGAATGAAGACCCACCTGTAGGTTTGGCAGCTAGCTTAAG 3960
Db |||||
3401 GTCTCCAGAAAAGGGGGAATGAAGACCCACCTGTAGGTTTGGCAGCTAGCTTAAG 3460
Qy |||||
3961 TAAACCCCATTTTGAAGGCATGG-AAAATAACATAACTGAGATAGAGAAGTTTCAATCAA 4019
Db |||||
3461 TAAACCCCATTTTGAAGGCATGGAATAATACATACTGAGATAGAGAAGTTTCAATCAA 3520
Qy |||||
4020 GGTTAGGAACAGA-GAGACAGCAATATGGGCCAAACAGGATATCTGTGGTAAGAGTT 4078
Db |||||
3521 GGTAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGAGTT 3580
Qy |||||
4079 CTTGCCCC- GCTCAGGGCCAAAGACAGTTGGAACAGGAGATATGGGCCAAACAGGATAT 4137
Db |||||
3581 CTTGCCCCGGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAACAGGATAT 3640
Qy |||||
4138 CTGTGTTAAGCAGTTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCCAGATGCGGT 4197

QY 6357 CCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGGAGCGCAGAAAGTGGTCTGCA 6416
 DB |||||
 QY 5469 CCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGCGGAGCGCAGAAAGTGGTCTGCA 5528
 DB |||||
 QY 6417 ACTTTATCGGCTCCATCAGTCTATTAAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCG 6476
 DB |||||
 QY 5529 ACTTTATCGGCTCCATCAGTCTATTAAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCG 5588
 DB |||||
 QY 6477 CCAGTTAATAGTTTGGCAACGTTGTCATTTGCTACAGGC-TGTTGGTGTACGCTCG 6535
 DB |||||
 QY 5589 CCAGTTAATAGTTTGGCAACGTTGTCATTTGCTACAGGCATCGTGGTGTACGCTCG 5648
 DB |||||
 QY 6536 TCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCC 6595
 DB |||||
 QY 5649 TCGTTTGGTATGGCTTCATTACAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCC 5708
 DB |||||
 QY 6596 CCATGTTGTCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTATCGTTGTCAAGTAAG 6655
 DB |||||
 QY 5709 CCATGTTGTCAAAAAGGGTTAGTCTCTTCGGTCTCCGATCGTATCGTTGTCAAGTAAG 5768
 DB |||||
 QY 6656 TTGGCCGAGTGTATCACTCATGTTATGGGAGCACTGCATAATTTCTTACTGTCTATG 6715
 DB |||||
 QY 5769 TTGGCCGAGTGTATCACTCATGTTATGGGAGCACTGCATAATTTCTTACTGTCTATG 5828
 DB |||||
 QY 6716 CCATCCGTAAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATTCAGAAATAG 6775
 DB |||||
 QY 5829 CCATCCGTAAAGATGCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCAATTCAGAAATAG 5888
 DB |||||
 QY 6776 TGTATCGGCGGACCGAGTTGCTCTTCCCGGGGTCAATACGGGATAAATACCGGCCACAT 6835
 DB |||||
 QY 5889 TGTATCGGCGGACCGAGTTGCTCTTCCCGGGGTCAATACGGGATAAATACCGGCCACAT 5948
 DB |||||
 QY 6836 AGCAGAACTTTAAAGTGTCTCATCTTGAAGACGTTCTTCGGGGGAAACCTCTCAAGG 6895
 DB |||||
 QY 5949 AGCAGAACTTTAAAGTGTCTCATCTTGAAGACGTTCTTCGGGGGAAACCTCTCAAGG 6008
 DB |||||
 QY 6896 ATCTTACCCTGTTGAGATCCAGTTCGATGTAACTCCACTCGTGCACTCACTGATCTTCA 6955
 DB |||||
 QY 6009 ATCTTACCCTGTTGAGATCCAGTTCGATGTAACTCCACTCGTGCACTCACTGATCTTCA 6068
 DB |||||
 QY 6956 GCATCTTTTACCTTCAACGAGTTCCTGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 7015
 DB |||||
 QY 6069 GCATCTTTTACCTTCAACGAGTTCCTGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 6128
 DB |||||
 QY 7016 AAAAGGGAATAAGGCGGACAGGAATGTTGAATACCTACTCTTCTTTTCAATAT 7075
 DB |||||
 QY 6129 AAAAGGGAATAAGGCGGACAGGAATGTTGAATACCTACTCTTCTTTTCAATAT 6188
 DB |||||
 QY 7076 TATTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAG 7135
 DB |||||
 QY 6189 TATTGAAGCATTTATCAGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAG 6248
 DB |||||
 QY 7136 AAAATAAACAATAGGGTTCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTAA 7195
 DB |||||
 QY 6249 AAAATAAACAATAGGGTTCGCGCAATTTCCCGGAAAGTGCCACCTGACGCTAA 6308
 DB |||||
 QY 7196 GAAACCATTTATCATGATTAACCTATAAATAGGGTATCACGAGGCCCTTTCGT 7255
 DB |||||
 QY 6309 GAAACCATTTATCATGATTAACCTATAAATAGGGTATCACGAGGCCCTTTCGT 6368
 DB |||||
 QY 7256 CT 7257
 DB ||
 QY 6369 CT 6370

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 25, 2005, 05:30:20 ; Search time 2422 Seconds
(without alignments)
18707.556 Million cell updates/sec

Title: US-10-006-771B-1
Perfect score: 7654
Sequence: 1 aagctgcagctgcaggt.....gtaaacagcggccagtgcc 7654

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7654	100.0	7654	8	ABX16565 Retrovira
2	7654	100.0	7654	10	ABX13168 Retrovira
3	3349	43.8	9320	6	ABX56664 Plasmid p
4	3307.8	43.2	6644	2	AAX33181 Base sequ
5	3244.4	42.4	6505	9	ADA12886 Murine MS
6	3223.6	42.1	5903	13	ADQ80674 Bcl-XL ex
7	3213	42.0	5782	6	ABX85887 DNA sequ
8	3183.2	41.6	7372	2	AAX33182 Base sequ
9	3139.6	41.0	6253	9	AAL57215 MIG retro
10	3130.2	40.9	7257	13	ADQ80673 Bcl-XL ex
11	3111	40.6	5364	2	AAT13393 Hybrid ve
12	3103.4	40.5	6444	6	AAL67595 Nucleotid
13	3096	40.4	6522	2	AAX90482 Plasmid r
14	3093	40.4	6221	3	AAX34935 Retrovira
15	3093	40.4	6221	4	AAX30944 Vector us
16	3092.4	40.4	6365	2	AAQ41173 Plasmid L
17	3083.4	40.3	5323	2	AAT13390 Hybrid ve
18	3080.2	40.2	5294	2	AAT13391 Hybrid ve
19	3077	40.2	5292	2	AAT13392 Hybrid ve
20	3062.2	40.0	6145	2	AAV83182 Intermedi

21	3062.2	40.0	6145	5	AAF85611	Aaf85611-ba
22	3050.2	39.9	6620	6	AD82653	Plasmid v
23	3006.8	39.3	6620	9	ADA12887	Murine le
24	2899.2	37.9	6141	2	AAX90481	Plasmid r
25	2814	36.8	6046	12	ADG83262	Clone pLX
26	2783.8	36.4	7797	2	AAX33180	Cowpox vi
27	2687.4	35.1	5874	2	AAX90484	Plasmid r
28	2683.4	35.1	5865	2	AAV04002	Retrovira
29	2675.4	35.0	6283	4	AAF83147	Complete
30	2672.4	34.9	7086	6	ABK12523	Tetracycl
31	2639	34.5	7165	2	AAX90483	Plasmid r
32	2634.8	34.4	5421	3	AAC68299	SV40/APPA
33	2634.8	34.4	6116	3	AAC68297	R15/APPA
34	2634.8	34.4	6708	3	AAC68295	R15/APPA
35	2634	34.4	6143	9	ADA12885	Murine re
36	2633.2	34.4	7560	4	AAC85599	Plasmid p
37	2633.2	34.4	7560	9	ADA09849	PiggyBac
38	2630.4	34.4	5715	3	AAX34936	Retroviru
39	2630.4	34.4	5715	4	AAF30945	Vector us
40	2627.6	34.3	4163	2	AAQ32349	Template
41	2624.4	34.3	5465	2	AAX20088	Plasmid p
42	2624	34.3	3309	6	AAK98929	Human pSG
43	2623.4	34.3	9164	2	AAX82259	Beta-doma
44	2623.4	34.3	11846	2	AAX82261	Factor VI
45	2623.4	34.3	12022	2	AAX82260	Factor VI

ALIGNMENTS

RESULT 1

ABX16565
ID ABX16565 standard; DNA; 7654 BP.

XX ABX16565;

DT 22-APR-2003 (first entry)

DE Retroviral vector expressing a chimaeric hMN14/T-cell receptor.

XX Retroviral vector; ds; T-cell receptor; hMN14; antibody; IgTCR;

KW cytosolic; dermatological; neuroprotective; immunostimulant; GD3;

KW ganglioside antigen; MB3.6; PSMA; tumour; 3D8; 4D4; 3E11;

KW prostate-specific membrane antigen; zeta signalling chain; human; cancer;

KW melanoma; neuroendocrine tumour; prostate cancer; small cell lung cancer;

KW mouse; CB8alpha hinge.

XX Homo sapiens.

OS Mus sp.

OS Retrovirus sp.

OS Synthetic.

OS Chimeric.

Key Location/Qualifiers
CDS 2428..3759

/*tag= a

/product= "IgTCR"

US2002132983-A1.

19-SEP-2002.

10-DEC-2001; 2001US-00006773.

30-NOV-2000; 2000US-0250087P.

30-NOV-2000; 2000US-0250089P.

(JUNG/) JUNGHANS R P.

Junghans RP;

WPI; 2003-208946/20.

P-ESDB; ABG74240.

XX New chimeric molecule useful in treating patients with disorders, such as
PT melanoma, neuroendocrine disorders, prostate and small cell lung cancer
PT comprise GD3 and/or PSMA binding domains of antibody.
XX
PS Disclosure; Page 3-7; 35pp; English.
XX

CC The invention relates to a chimeric molecule comprising the GD3
CC (ganglioside antigen) binding domain of antibody MB3.6, with any of 3
CC variable gene sequences, or the PSMA (prostate-specific membrane antigen)
CC binding domain of antibody 3D8, 4D4 and 3E11, with variable gene
CC sequences, the zeta signalling chain of the T cell receptor and an
CC intervening CD8alpha hinge in which cysteine residues have been mutated.
CC The chimeric molecules expressed in T cells or NK cells or other
CC effector cells are useful in treating patients with cancers expressing
CC the GD3 (MB3.6 derivatives) or PSMA antigen (3D8, 4D4, 3E11 derivatives),
CC and/or together with each other or with heterologous constructs to engage
CC additional stimulatory and functional properties of the effector cells to
CC enhance the antitumour therapeutic efficacy (claimed). They are
CC particularly useful in disorders including melanoma, neuroendocrine
CC tumours and prostate and small cell lung cancer. The present sequence is
CC a retroviral vector encoding the hMN14 antibody (specific to CEA antigen)
CC in a fusion protein with the modified CD8alpha hinge and the T-cell
CC receptor zeta chain (IgTCR). The hMN14 antibody coding region is replaced
CC with the MB3.6, 3D8, 4D4 or 3E11 genes of the invention
XX

SQ Sequence 7654 BP; 1832 A; 2067 C; 1953 G; 1802 T; 0 U; 0 Other;

Query Match 100.0%; Score 7654; DB 8; Length 7654;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTTGCATGCTGAGTGGAGGCGGTTAAACCGTTTAAACGGTTTAAATAAATCAATTAATTT 60
DB 1 AAGCTTGCATGCTGAGTGGAGGCGGTTAAACCGTTTAAACGGTTTAAATAAATCAATTAATTT 60
QY 61 GCAGCCGACAGTGAAGAGACCGTTAAACCGTTTAAACGGTTTAAATAAATCAATTAATTT 120
DB 61 GCAGCCGACAGTGAAGAGACCGTTAAACCGTTTAAACGGTTTAAATAAATCAATTAATTT 120
QY 121 AGAGTCATTTCTTTGGTAGGAAGTACATTTGGCAGTAAAGAGGCCCAAGCAATCTGTG 180
DB 121 AGAGTCATTTCTTTGGTAGGAAGTACATTTGGCAGTAAAGAGGCCCAAGCAATCTGTG 180
QY 181 GAAAGCCAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCGCGTGTACCTAAGGTTT 240
DB 181 GAAAGCCAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCGCGTGTACCTAAGGTTT 240
QY 241 CTTAATTTGTGGTTTCTAAATCTTCCAGAGGTTTGTCTCAFTTCACTTCCACTTCGGTG 300
DB 241 CTTAATTTGTGGTTTCTAAATCTTCCAGAGGTTTGTCTCAFTTCACTTCCACTTCGGTG 300
QY 301 CACAATACCTTGGACGGGATTTACTGTCTTAGCATCTATCGTGGCCCTTCGATTGAGGC 360
DB 301 CACAATACCTTGGACGGGATTTACTGTCTTAGCATCTATCGTGGCCCTTCGATTGAGGC 360
QY 361 TGAACCTGAGGCCACTCTTCTAGCTTTAGAGGAGCACAAGCACAAGAGAGGCTGA 420
DB 361 TGAACCTGAGGCCACTCTTCTAGCTTTAGAGGAGCACAAGCACAAGAGAGGCTGA 420
QY 421 CCGGCGAGACCTGTGGGCAATTTTAAACAGGCGCTCTCGGTCTGTGGGAGGCGAGCTTA 480
DB 421 CCGGCGAGACCTGTGGGCAATTTTAAACAGGCGCTCTCGGTCTGTGGGAGGCGAGCTTA 480
QY 481 CATAAGGTGCAATTAAGAATATAAATAAGCCCATATCAATTTGTCTATCTTTTAA 540
DB 481 CATAAGGTGCAATTAAGAATATAAATAAGCCCATATCAATTTGTCTATCTTTTAA 540
QY 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATT 600
DB 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATT 600
QY 601 TTGCAAGGATGGAATAATACATACTGAGAATAGAGAAAGTTTCAGATCAAGGTTAGGAACA 660

DB 601 TTGCAAGGATGGAATAATACATACTGAGAATAGAGAAAGTTTCAGATCAAGGTTAGGAACA 660
QY 661 GAGAGACAGCAGAAATATGGGCCAAACAGAGATATCTGTGTAAAGAGTTCTCTCGCCCGCTC 720
DB 661 GAGAGACAGCAGAAATATGGGCCAAACAGAGATATCTGTGTAAAGAGTTCTCTCGCCCGCTC 720
QY 721 AGGGCCAAAGAACAGTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAAAGAG 780
DB 721 AGGGCCAAAGAACAGTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAAAGAG 780
QY 781 TTCTCTGCCCGGCTCAGGCGCCAAAGAACAGAGTGGTCCCCAGATGCGGTCCCGCCCTCAGCA 840
DB 781 TTCTCTGCCCGGCTCAGGCGCCAAAGAACAGAGTGGTCCCCAGATGCGGTCCCGCCCTCAGCA 840
QY 841 GTTCTTAGAGAACCAATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGC 900
DB 841 GTTCTTAGAGAACCAATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGC 900
QY 901 CTTATTTGAACTAAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCG 960
DB 901 CTTATTTGAACTAAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCG 960
QY 961 AGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGT 1020
DB 961 AGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGT 1020
QY 1021 GCGCCGGGTACCCGTATTTCCCAATAAAGCCTTCTGTTGTCATCCGAATCGTGGACTC 1080
DB 1021 GCGCCGGGTACCCGTATTTCCCAATAAAGCCTTCTGTTGTCATCCGAATCGTGGACTC 1080
QY 1081 GCTGATCTTGGAGGGTCTCTCAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1140
DB 1081 GCTGATCTTGGAGGGTCTCTCAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1140
QY 1141 TGGAGGTTTCCACCCGAGATTTGGAGACCCCTGCCCCAGGACCAACCCCGCCCGGGA 1200
DB 1141 TGGAGGTTTCCACCCGAGATTTGGAGACCCCTGCCCCAGGACCAACCCCGCCCGGGA 1200
QY 1201 GGTAAAGTGGCCAGCAACTTATCTGTGTCGTCGATTTGTCTAGTGTCTATGACTGATTT 1260
DB 1201 GGTAAAGTGGCCAGCAACTTATCTGTGTCGTCGATTTGTCTAGTGTCTATGACTGATTT 1260
QY 1261 TATGCGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGAGCCCGTGGTG 1320
DB 1261 TATGCGCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGCGGAGCCCGTGGTG 1320
QY 1321 AACTGACAGTTTCCGGAACACCCGCGCCCAACCCCTGGGAGAGCTCCAGGAGCTTCGGGG 1380
DB 1321 AACTGACAGTTTCCGGAACACCCGCGCCCAACCCCTGGGAGAGCTCCAGGAGCTTCGGGG 1380
QY 1381 CCGTTTTTGTGGCCGACCTGAGTCTTAAATCCCGATCGTTTGGAGCTCTTTGGTGCAC 1440
DB 1381 CCGTTTTTGTGGCCGACCTGAGTCTTAAATCCCGATCGTTTGGAGCTCTTTGGTGCAC 1440
QY 1441 CCCCCCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGC 1500
DB 1441 CCCCCCTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTTCCCGC 1500
QY 1501 CTCCTGCTGAATTTTGTCTTGGGACCCGAGCCGCGCGCTCTTGTCTGCT 1560
DB 1501 CTCCTGCTGAATTTTGTCTTGGGACCCGAGCCGCGCGCTCTTGTCTGCT 1560
QY 1561 GCAGCATCGTTCTGTGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1561 GCAGCATCGTTCTGTGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1621 GCCCGGCTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGTGTACCTGGAAGATGTCG 1680
DB 1621 GCCCGGCTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGTGTACCTGGAAGATGTCG 1680
QY 1681 AGCGGATCGCTCACCAACCAAGTGGTAGATGTCAGAGAGAGAGTGGGTTACCTTCTGCT 1740

Db 1681 AGCGATCGCTCACACCAGTCGGTAGATGTCAAGAAAGAGAGGTTGGGTTACCTTCTGCT 1740
Qy 1741 CTGCAGAAATGCCAACCTTTAACTCGGATGCGCGAGAGCGCACTTTTAAACCGAGACC 1800
Db 1741 CTGCAGAAATGCCAACCTTTAACTCGGATGCGCGAGAGCGCACTTTTAAACCGAGACC 1800
Qy 1801 TCATCACCAGGTTAAGATCAAGGTCCTTTTACCTGGSCCGCATGGACACCCAGACGAG 1860
Db 1801 TCATCACCAGGTTAAGATCAAGGTCCTTTTACCTGGSCCGCATGGACACCCAGACGAG 1860
Qy 1861 TCCCTCATCTGTCAGCTCGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGCCCT 1920
Db 1861 TCCCTCATCTGTCAGCTCGGAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAGGCCCT 1920
Qy 1921 TTGTACACCTTAAGCTCCGCTCTCTCTTCATCGCCGCCCTCTCTCCCTCTGAAC 1980
Db 1921 TTGTACACCTTAAGCTCCGCTCTCTCTTCATCGCCGCCCTCTCTCCCTCTGAAC 1980
Qy 1981 CTCTCTGTTGACCCCGCTCGATCCTCTCTTTATCGAGCCCTCACTCTCTTAGGCG 2040
Db 1981 CTCTCTGTTGACCCCGCTCGATCCTCTCTTTATCGAGCCCTCACTCTCTTAGGCG 2040
Qy 2041 CCCCATATGSCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAACCTTCCCTG 2100
Db 2041 CCCCATATGSCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAACCTTCCCTG 2100
Qy 2101 ACCCTGACATGACAAAGGTTACTTAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160
Db 2101 ACCCTGACATGACAAAGGTTACTTAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160
Qy 2161 CTTAGTCAGCAACGAACTCTGGGAGCTCTGGGAGCTTCAAGAACAACTTGAAC 2220
Db 2161 CTTAGTCAGCAACGAACTCTGGGAGCTCTGGGAGCTTCAAGAACAACTTGAAC 2220
Qy 2221 ACCGCTGTACTCTACCTTACCGAGTCGGGACACAGTGTGGTCCGCCACACAGAC 2280
Db 2221 ACCGCTGTACTCTACCTTACCGAGTCGGGAGCTTCAAGAACAACTTGAAC 2280
Qy 2281 TAAGAACTAGAACCTCTGCTGAAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 2340
Db 2281 TAAGAACTAGAACCTCTGCTGAAAGGACCTTACACAGTCTCTGTGACACCCCGACCGC 2340
Qy 2341 CCTCAAAGTAGACGGCATCGAGCTTGATATACACGCGCCCACTGAAAGGTCGCGACCC 2400
Db 2341 CCTCAAAGTAGACGGCATCGAGCTTGATATACACGCGCCCACTGAAAGGTCGCGACCC 2400
Qy 2401 CGGGGTTGACCATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTGTA 2460
Db 2401 CGGGGTTGACCATCTCTAGACTGCCATGGATGGAGCTGTATCATCTCTTCTTGTA 2460
Qy 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTTACG 2520
Db 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTTACG 2520
Qy 2521 GCGAGCGTGGTGACAGAGTCAACATCACTGTAAAGGCGAGTCAAGGATGCGGACTTCT 2580
Db 2521 GCGAGCGTGGTGACAGAGTCAACATCACTGTAAAGGCGAGTCAAGGATGCGGACTTCT 2580
Qy 2581 GTAGCTTGGTACAGCAGAACCGAGTAAAGCTCCAAAGCTGCTGATCTACTGACATCC 2640
Db 2581 GTAGCTTGGTACAGCAGAACCGAGTAAAGCTCCAAAGCTGCTGATCTACTGACATCC 2640
Qy 2641 ACCGGCACACTGGTGTCCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGATTCACC 2700
Db 2641 ACCGGCACACTGGTGTCCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGATTCACC 2700
Qy 2701 TTCAACCATCAGCAGCTCCAGCAGAGGACATCGCACTACTACTGCGCAATATAGC 2760
Db 2701 TTCAACCATCAGCAGCTCCAGCAGAGGACATCGCACTACTACTGCGCAATATAGC 2760
Qy 2761 CTCTATCGGTCGTCGCGCAAGGACCAAGGTGAAATCAACAGGTCGTCAGGATCG 2820
Db 2761 CTCTATCGGTCGTCGCGCAAGGACCAAGGTGAAATCAACAGGTCGTCAGGATCG 2820

Qy 2821 GGTGGATCCGCTCTGCTGGCTCAGGATCGAGGTCCAACCTGGTGAGAGCGGTGAGGT 2880
Db 2821 GGTGGATCCGCTCTGCTGGCTCAGGATCGAGGTCCAACCTGGTGAGAGCGGTGAGGT 2880
Qy 2881 GTTGTGCAACTGCGCGCTCCCTGGCTGCTGCTCCGATCTGCGCTTCTGATTTTACC 2940
Db 2881 GTTGTGCAACTGCGCGCTCCCTGGCTGCTGCTCCGATCTGCGCTTCTGATTTTACC 2940
Qy 2941 ACATATTGGATGAGTGGGTGAGACAGGACCTCGAAAAGGTCTTGAAGTGGAGAA 3000
Db 2941 ACATATTGGATGAGTGGGTGAGACAGGACCTCGAAAAGGTCTTGAAGTGGAGAA 3000
Qy 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTACAATA 3060
Db 3001 ATTCAATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTACAATA 3060
Qy 3061 TCGGAGACAAACGCAAGAACACATTGCTCTGCAAAATGGACAGCTGAGACCCGAGAC 3120
Db 3061 TCGGAGACAAACGCAAGAACACATTGCTCTGCAAAATGGACAGCTGAGACCCGAGAC 3120
Qy 3121 ACCGGGCTCTATTTTGTGCAAGCTTAACTTCTGCGCTTCCCTGCTTGTATTTGGGCG 3180
Db 3121 ACCGGGCTCTATTTTGTGCAAGCTTAACTTCTGCGCTTCCCTGCTTGTATTTGGGCG 3180
Qy 3181 CAAGGACCCCGGTCACTGCTTCCAGTGTCTAAGCCCAACACAGCCAGCCGCGACCA 3240
Db 3181 CAAGGACCCCGGTCACTGCTTCCAGTGTCTAAGCCCAACACAGCCAGCCGCGACCA 3240
Qy 3241 CCAACACGCGCGCCCACTGCTGCGAGCCCTGCTGCTGCGCCAGAGCGGCTCGG 3300
Db 3241 CCAACACGCGCGCCCACTGCTGCGAGCCCTGCTGCTGCGCCAGAGCGGCTCGG 3300
Qy 3301 CCAGCGCGGGGCGCAGTGCACACGAGGGGCTGACCTTCCGCTTGATCCCAAACTC 3360
Db 3301 CCAGCGCGGGGCGCAGTGCACACGAGGGGCTGACCTTCCGCTTGATCCCAAACTC 3360
Qy 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATGCTGTCAFTTCTCACTGCTTGTCTCG 3420
Db 3361 TGCTACCTGCTGGATGGAATCTCTTCACTATGCTGTCAFTTCTCACTGCTTGTCTCG 3420
Qy 3421 AGAGTGAAGTTTACAGCAGGCGCAGAGCCCGCCGCTACAGAGGGGCGAGACGAGTC 3480
Db 3421 AGAGTGAAGTTTACAGCAGGCGCAGAGCCCGCCGCTACAGAGGGGCGAGACGAGTC 3480
Qy 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGCG 3540
Db 3481 TATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGCG 3540
Qy 3541 CGGACCCCTGAGATGGGGGAAAGCCGAGAAAGAAACCCCTCAGGAAGGCTGTACAAAT 3600
Db 3541 CGGACCCCTGAGATGGGGGAAAGCCGAGAAAGAAACCCCTCAGGAAGGCTGTACAAAT 3600
Qy 3601 GAATCGCAGAAAGATAGATGGCGGAGCCCTACAGTGAAGTGGGATGAAAGGCGAGCGC 3660
Db 3601 GAATCGCAGAAAGATAGATGGCGGAGCCCTACAGTGAAGTGGGATGAAAGGCGAGCGC 3660
Qy 3661 CGGAGGGGCAAGGGCAGATGCGCTTTACAGGGTCTCAGTACAGCACCAGAGACACC 3720
Db 3661 CGGAGGGGCAAGGGCAGATGCGCTTTACAGGGTCTCAGTACAGCACCAGAGACACC 3720
Qy 3721 TACGACCCCTTCACTGACAGGCGCTGCGCCCTCGCTAACTCGACGCGGCGCGGATCCG 3780
Db 3721 TACGACCCCTTCACTGACAGGCGCTGCGCCCTCGCTAACTCGACGCGGCGCGGATCCG 3780
Qy 3781 GATTAGTCCAAATTTGTTTAAAGACAGAGTATCAGTGGTCCAGGCTCTAGTTTGTACTAAC 3840
Db 3781 GATTAGTCCAAATTTGTTTAAAGACAGAGTATCAGTGGTCCAGGCTCTAGTTTGTACTAAC 3840
Qy 3841 AATATCACAGCTGAAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900
Db 3841 AATATCACAGCTGAAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900

QY 3901 GTCTCAGAAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960
DB 3901 GTCTCAGAAAAAGGGGGAATGAAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960
QY 3961 TAACGCCATTTTGCAGGCAATGGAATAATACATAACTTGAGATAGAGAAGTTGAGATCAAG 4020
DB 3961 TAACGCCATTTTGCAGGCAATGGAATAATACATAACTTGAGATAGAGAAGTTGAGATCAAG 4020
QY 4021 GTTAGGAACAGAGAGACAGAGAAATGGGCCAAACAGGATATCTGTGTAAAGCAGTTC 4080
DB 4021 GTTAGGAACAGAGAGACAGAGAAATGGGCCAAACAGGATATCTGTGTAAAGCAGTTC 4080
QY 4081 TGCCCCGCTCAGGGCCAAAGACAGTTGGAAACAGGAGATATGGGCCAAACAGGATATCTG 4140
DB 4081 TGCCCCGCTCAGGGCCAAAGACAGTTGGAAACAGGAGATATGGGCCAAACAGGATATCTG 4140
QY 4141 TGTAAAGCAGTTCCTGCCCGGCTCAGGGCCAAAGACAGATGGTTCCTCAGATGCGGTCCC 4200
DB 4141 TGTAAAGCAGTTCCTGCCCGGCTCAGGGCCAAAGACAGATGGTTCCTCAGATGCGGTCCC 4200
QY 4201 GCCCTCAGCAGTTTCTAGAGACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAT 4260
DB 4201 GCCCTCAGCAGTTTCTAGAGACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAT 4260
QY 4261 GACCTGTGCTTATTTGAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCTT 4320
DB 4261 GACCTGTGCTTATTTGAACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCTT 4320
QY 4321 CTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGGCGCAGTCTCTCCGA 4380
DB 4321 CTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGGCGCAGTCTCTCCGA 4380
QY 4381 TAGACTGCGTCCGCGGCTACCGGTCTCTGAGTGATTAACAAACCTCTTGAGTTGATCCGACT 4440
DB 4381 TAGACTGCGTCCGCGGCTACCGGTCTCTGAGTGATTAACAAACCTCTTGAGTTGATCCGACT 4440
QY 4441 CGTGTGCTCGCTTCTTGGGAGGCTCTCTGAGTGATTAACAAACCTCTTGAGTTGATCCGACT 4500
DB 4441 CGTGTGCTCGCTTCTTGGGAGGCTCTCTGAGTGATTAACAAACCTCTTGAGTTGATCCGACT 4500
QY 4501 CTTTTCAGTTTCTCCCACTACACAGGCTCACTAAATTCCTGATGTCGCGAGGACTC 4560
DB 4501 CTTTTCAGTTTCTCCCACTACACAGGCTCACTAAATTCCTGATGTCGCGAGGACTC 4560
QY 4561 CGTCAGCCGCTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGACT 4620
DB 4561 CGTCAGCCGCTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGACT 4620
QY 4621 ACATCTGACTCTCGGCTTTATAAAGAATGTTGAAGGCTCTGTGGAATATCTGCGACA 4680
DB 4621 ACATCTGACTCTCGGCTTTATAAAGAATGTTGAAGGCTCTGTGGAATATCTGCGACA 4680
QY 4681 CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTGCTA 4740
DB 4681 CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTGCTA 4740
QY 4741 TTTTGTGCTCTCTCGTATTTTAAATTTCTAGTTTGCATCCCTTCTGAGAGCACGCG 4800
DB 4741 TTTTGTGCTCTCTCGTATTTTAAATTTCTAGTTTGCATCCCTTCTGAGAGCACGCG 4800
QY 4801 GATTGACAGATAGTTAATACTCTGAGGCGAGGCTCTGTGAAAGGTTGCTGGGCTCAG 4860
DB 4801 GATTGACAGATAGTTAATACTCTGAGGCGAGGCTCTGTGAAAGGTTGCTGGGCTCAG 4860
QY 4861 TGTGAGATTTGCCATAAAGGGTCTGCGCCCTGTGTACAGACAGATCGGAATCTAGA 4920
DB 4861 TGTGAGATTTGCCATAAAGGGTCTGCGCCCTGTGTACAGACAGATCGGAATCTAGA 4920
QY 4921 GTGCATCTCAGAGTCCCGGCTCTGATCTCAGGCGATCTTTGCTCCTAGAG 4980
DB 4921 GTGCATCTCAGAGTCCCGGCTCTGATCTCAGGCGATCTTTGCTCCTAGAG 4980
QY 4981 ATCCTCTACGCCGACGCATCGTGGCGGGGTACCGAGCTCGAATTCGTAATCATGTGTCAT 5040

DB 4981 ATCCTCTACGCCGACGCATCGTGGCGGGGTACCGAGCTCGAATTCGTAATCATGTGTCAT 5040
QY 5041 AGCTGTTCCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACAACTACGAGCGGAA 5100
DB 5041 AGCTGTTCCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACAACTACGAGCGGAA 5100
QY 5101 GCATAAAGTGTAAAGCTTGGGGTGCTAATAGTGAAGCTAACTCACATTAATTTGGCTTGC 5160
DB 5101 GCATAAAGTGTAAAGCTTGGGGTGCTAATAGTGAAGCTAACTCACATTAATTTGGCTTGC 5160
QY 5161 GCTCACTGCTGCTTTCAGTTCGGGAAACCTGTCTGTGCGAGCTGCAATTAATGAATCGGCC 5220
DB 5161 GCTCACTGCTGCTTTCAGTTCGGGAAACCTGTCTGTGCGAGCTGCAATTAATGAATCGGCC 5220
QY 5221 AACGCGGGGAGAGCGGTTTGCATATTTGGCGCTCTTCCGCTTCTCGCTCAGCTGACT 5280
DB 5221 AACGCGGGGAGAGCGGTTTGCATATTTGGCGCTCTTCCGCTTCTCGCTCAGCTGACT 5280
QY 5281 CGCTCGCTCGCTTCTGCTGCGGCTGCGGAGCGGTTATCAGCTCACTCAAAAGCGGTAATAC 5340
DB 5281 CGCTCGCTCGCTTCTGCTGCGGCTGCGGAGCGGTTATCAGCTCACTCAAAAGCGGTAATAC 5340
QY 5341 GGTATTCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA 5400
DB 5341 GGTATTCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA 5400
QY 5401 AGCCAGGAAACCGTAAAAAGGCGGTTGCTGCGGTTTCCATAGGCTCCGCCCCCTG 5460
DB 5401 AGCCAGGAAACCGTAAAAAGGCGGTTGCTGCGGTTTCCATAGGCTCCGCCCCCTG 5460
QY 5461 ACAGCATCACAAAAATCGACGCTCAAGTCAAGGTGGCGAAACCGCAGCAGACTATAAA 5520
DB 5461 ACAGCATCACAAAAATCGACGCTCAAGTCAAGGTGGCGAAACCGCAGCAGACTATAAA 5520
QY 5521 GATACAGCGGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCTGTTCGACCCCTGCGCC 5580
DB 5521 GATACAGCGGTTTCCCTCGGAAAGCTCCCTCGTGGCTCTCTGTTCGACCCCTGCGCC 5580
QY 5581 TTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTAC 5640
DB 5581 TTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGCGCTTCTCATAGCTAC 5640
QY 5641 GCTGTAGGTATCTCAGTTTCGCTGAGTGTGCTGCTCAAGCTGGGCTGTGTGACAGAAC 5700
DB 5641 GCTGTAGGTATCTCAGTTTCGCTGAGTGTGCTGCTCAAGCTGGGCTGTGTGACAGAAC 5700
QY 5701 CCCCCCTTACGCCGACCGCTGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCG 5760
DB 5701 CCCCCCTTACGCCGACCGCTGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCG 5760
QY 5761 TAAGACACGACTTATCGCCACTGGGAGGACCACTGTTAAAGGATTTAGCAGGCGAGT 5820
DB 5761 TAAGACACGACTTATCGCCACTGGGAGGACCACTGTTAAAGGATTTAGCAGGCGAGT 5820
QY 5821 ATGTAGGCGGCTACAGAGTTCTTGAAGTGTGCTTAACTACGCTACCTAGAGGA 5880
DB 5821 ATGTAGGCGGCTACAGAGTTCTTGAAGTGTGCTTAACTACGCTACCTAGAGGA 5880
QY 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAAGGCTTACCTTCGGAAGGATTTGGTAGCT 5940
DB 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAAGGCTTACCTTCGGAAGGATTTGGTAGCT 5940
QY 5941 CTTGATTCGGGCAAAACCAACCGCTGTTAGGCTGTTTGTGTTGCAAGCAGCAGA 6000
DB 5941 CTTGATTCGGGCAAAACCAACCGCTGTTAGGCTGTTTGTGTTGCAAGCAGCAGA 6000
QY 6001 TTACGCGCAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTGACG 6060
DB 6001 TTACGCGCAAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTGACG 6060
QY 6061 CTTAGTGGAAACCACTCACTTAAAGGATTTGGTCTAGAGATTAATCAAAAGGATCT 6120
DB 6061 CTTAGTGGAAACCACTCACTTAAAGGATTTGGTCTAGAGATTAATCAAAAGGATCT 6120

Db 6061 CTCAGTGGAAAGAACTCAGCTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 6120
Qy 6121 TCACCTAGATCCTTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATAGT 6180
Db 6121 TCACCTAGATCCTTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATAGT 6180
Qy 6181 AAATCTGCTGACAGTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATAGT 6240
Db 6181 AAATCTGCTGACAGTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATAGT 6240
Qy 6241 TATTTCTGCTCATCATAGTTCCTGACTCCCGCTGCTAGATAAATACGATACGGAGG 6300
Db 6241 TATTTCTGCTCATCATAGTTCCTGACTCCCGCTGCTAGATAAATACGATACGGAGG 6300
Qy 6301 CTTTACCATCTGCGCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360
Db 6301 CTTTACCATCTGCGCCAGTCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6360
Qy 6361 ATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTCTGCACTT 6420
Db 6361 ATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTCTGCACTT 6420
Qy 6421 TATCGGCTCCATCCAGTCTTAAATTTGTCGCGGAAGCTAGAGTAAGTGTTCGCCAG 6480
Db 6421 TATCGGCTCCATCCAGTCTTAAATTTGTCGCGGAAGCTAGAGTAAGTGTTCGCCAG 6480
Qy 6481 TTAATAGTTTCGCAACGTTGTCGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG 6540
Db 6481 TTAATAGTTTCGCAACGTTGTCGCAATGATACCGCGAGTGGTGTTCGCCAGTCTCGT 6540
Qy 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCGCAT 6600
Db 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGCGAGTTACATGATCCCGCAT 6600
Qy 6601 GTTGTGCAAAAGCGGTAGCTCTTCGCTCCGATCTGTCGAGTAAGTGTGGC 6660
Db 6601 GTTGTGCAAAAGCGGTAGCTCTTCGCTCCGATCTGTCGAGTAAGTGTGGC 6660
Qy 6661 CGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTTCTTACTGTATGCCATC 6720
Db 6661 CGCAGTGTATCACTCATGTTTATGGCAGCACTGCATAATTTCTTACTGTATGCCATC 6720
Qy 6721 CGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTAT 6780
Db 6721 CGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGAAATAGTGTAT 6780
Qy 6781 CGCGGACCGAGTGTCTTCCCGGCTCAATACGGGATAATACCGCGCCACATAGCAG 6840
Db 6781 CGCGGACCGAGTGTCTTCCCGGCTCAATACGGGATAATACCGCGCCACATAGCAG 6840
Qy 6841 AACTTTAAAGTGTCTCATCATTTGGAACAGTTCTTCGGGGGAAACCTCTCAAGATCTT 6900
Db 6841 AACTTTAAAGTGTCTCATCATTTGGAACAGTTCTTCGGGGGAAACCTCTCAAGATCTT 6900
Qy 6901 ACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCAACCACTGATCTTCAGCATC 6960
Db 6901 ACCGCTGTGAGATCCAGTTCGATTAACCCACTCGTGCAACCACTGATCTTCAGCATC 6960
Qy 6961 TTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATGCGGCAAAAAA 7020
Db 6961 TTTTACTTTTCCAGCGTTTCTGGGTGAGCAAAACAGGAGGCAAAATGCGGCAAAAAA 7020
Qy 7021 GGGAAATAGGCGGACACAGGAATGTTGAATACTCATACTCTTCTTTTCAATATTATG 7080
Db 7021 GGGAAATAGGCGGACACAGGAATGTTGAATACTCATACTCTTCTTTTCAATATTATG 7080
Qy 7081 AAGCAATTTATCAGGGTTATGTCATCAGCGGATACATATTGAATGATTTTGAAGAAA 7140
Db 7081 AAGCAATTTATCAGGGTTATGTCATCAGCGGATACATATTGAATGATTTTGAAGAAA 7140
Qy 7141 TAAACAAATAGGGTTTCCGCGACATTTCCCGGAAAGTGCACCTGACGCTTAAGAAC 7200
Db 7141 TAAACAAATAGGGTTTCCGCGACATTTCCCGGAAAGTGCACCTGACGCTTAAGAAC 7200

Qy 7201 CATTATTATCATGATTAACCTATAAAATAGCGGTATACAGAGGCCCTTTTCGTCTGCG 7260
Db 7201 CATTATTATCATGATTAACCTATAAAATAGCGGTATACAGAGGCCCTTTTCGTCTGCG 7260
Qy 7261 GCGTTTCGCTGATGACCGTGAACCTCTGACACATCGAGCTCCCGGAGCGGTACACAG 7320
Db 7261 GCGTTTCGCTGATGACCGTGAACCTCTGACACATCGAGCTCCCGGAGCGGTACACAG 7320
Qy 7321 TTGTCTCTAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGG 7380
Db 7321 TTGTCTCTAAGCGGATCCCGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGG 7380
Qy 7381 CGGCTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCA 7440
Db 7381 CGGCTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCA 7440
Qy 7441 TATGCGGTGTCAATATCCGACAGATCGGTAAGGAGAAATACCGCATCAGCGGCCATTC 7500
Db 7441 TATGCGGTGTCAATATCCGACAGATCGGTAAGGAGAAATACCGCATCAGCGGCCATTC 7500
Qy 7501 GCCATTGAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACG 7560
Db 7501 GCCATTGAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGCTCTTCGCTATTACG 7560
Qy 7561 CCAGCTGCGGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTC 7620
Db 7561 CCAGCTGCGGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTC 7620
Qy 7621 CCAGTCAAGGCTGTAAACGAGCGCGCATGTC 7654
Db 7621 CCAGTCAAGGCTGTAAACGAGCGCGCATGTC 7654

RESULT 2

ABX13168
ID ABX13168 standard; DNA; 7654 BP.

XX ABX13168;

DT 13-MAY-2003 (first entry)

XX Retroviral vector encoding a humanised anti-CEA antibody, hMN14.

CEA; carcinoembryonic antigen; ds; IgTCR; T-cell receptor; cancer;
tumour; colorectal cancer; breast cancer; lung cancer; hMN14; cytostatic;
mouse; human; zeta signalling chain; CD8alpha hinge; humanised antibody;
retroviral vector.

OS Retrovirus sp.

OS Homo sapiens.

OS Mus sp.

OS Synthetic.

OS Chimeric.

XX Key Location/Qualifiers

FH CDS 2428..3759

FT /*tag= a

FT /product= "Chimaeric IgTCR protein"

XX US2002165360-A1.

XX 07-NOV-2002.

XX 10-DEC-2001; 2001US-00006771.

XX 30-NOV-2000; 2000US-0250087P.

XX 30-NOV-2000; 2000US-0250090P.

XX (JUNG/) JUNGHANS R P.

XX Junghans RP;

PI

XX

DR WPI; 2003-298705/29.
 DR P-PSDB; ABG76488.
 XX New chimeric molecule from humanized antibody against carcinoembryonic
 PT antigen and having signaling molecules of T cells and other effector
 PT cells, useful for the treatment of colorectal, breast and lung cancers.
 XX Disclosure; Page 3-7; 20pp; English.
 XX The invention relates to a chimaeric molecule comprising the
 CC carcinoembryonic antigen (CEA) binding domain of humanised antibody MN14
 CC as a single chain antibody with a (GGSGS)3 linker, the zeta signaling
 CC chain of the T cell receptor (TCR) and an intervening CD8alpha hinge in
 CC which the cysteine residues have been mutated, with the IgTCR molecule
 CC occupying nucleotides 2426-3766 of the retroviral vector sequence.
 CC appearing as ABX1368. The new chimaeric molecule expressed in T cells,
 CC NK (not defined) or other effector cells are useful in treating patients
 CC with cancers expressing the CEA antigen, together with other or with
 CC heterologous constructs to engage additional stimulatory and functional
 CC properties of the effector cells to enhance the anti-tumour therapeutic
 CC efficacy. The cancer disorder includes colorectal, breast and lung
 CC cancers. The present sequence encodes the chimaeric molecule of the
 CC invention. Note: Claim 1 refers to a sequence shown in figure 3. Figure 3
 CC does not contain a sequence, merely a diagram of the present sequence
 CC which only appears in the sequence listing
 XX
 SQ Sequence 7654 BP; 1832 A; 2067 C; 1953 G; 1802 T; 0 U; 0 Other;
 Query Match 100.0%; Score 7654; DB 10; Length 7654;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 7654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGCTTGCATGCTCGAGTCCGACTCTAGGCATATAAGAAACATAACTAACCAAGCT 60
 DB 1 AAGCTTGCATGCTCGAGTCCGACTCTAGGCATATAAGAAACATAACTAACCAAGCT 60
 QY 61 GCAGCCGAGACAGTGAAGAACCGTTAAACGGTTGTTTAAATAAACTGAATATATT 120
 DB 61 GCAGCCGAGACAGTGAAGAACCGTTAAACGGTTGTTTAAATAAACTGAATATATT 120
 QY 121 AGAGTCATTTCTTGGTAGAAGAGTACATTTGGCACGTAAAGGAGCCCAAAGCAATCTGTG 180
 DB 121 AGAGTCATTTCTTGGTAGAAGAGTACATTTGGCACGTAAAGGAGCCCAAAGCAATCTGTG 180
 QY 181 GAAAGCCAGGCTGGAGCCCGACAGTTTGCATCCCTCTCGCGTGTACCTAAGGGTTT 240
 DB 181 GAAAGCCAGGCTGGAGCCCGACAGTTTGCATCCCTCTCGCGTGTACCTAAGGGTTT 240
 QY 241 CTTAAATTGTGGTTTCTAAATCTTCCAGAGGTTTGTCTCATTTCCACTTCCGTTG 300
 DB 241 CTTAAATTGTGGTTTCTAAATCTTCCAGAGGTTTGTCTCATTTCCACTTCCGTTG 300
 QY 301 CACAATACCTTGGACCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATGAGGC 360
 DB 301 CACAATACCTTGGACCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATGAGGC 360
 QY 361 TGAACCTGAGGCCCTCTTTCAGCTTGTAAAGGAGCACAAGACCAAGAGAGCTGA 420
 DB 361 TGAACCTGAGGCCCTCTTTCAGCTTGTAAAGGAGCACAAGACCAAGAGAGCTGA 420
 QY 421 CCCGGCAGACCTGTGGCATTTTAAACAAGGGCTCTCTGGGTCTGTGGGAGGAGGCTTA 480
 DB 421 CCCGGCAGACCTGTGGCATTTTAAACAAGGGCTCTCTGGGTCTGTGGGAGGAGGCTTA 480
 QY 481 CATAAGGTGCAAAATTAGAAATATAAATAAAGCCCATATCAATTTGTCACTTTTTTTA 540
 DB 481 CATAAGGTGCAAAATTAGAAATATAAATAAAGCCCATATCAATTTGTCACTTTTTTTA 540
 QY 541 AGCTCAAGTTTGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTACGCCATT 600
 DB 541 AGCTCAAGTTTGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTACGCCATT 600
 QY 601 TTGCAAGGCAATGGAATAATACATAACTGAGAAATAGAGAAGTTTACAGATCAAGGTTAGGAACA 660

601 TTGCAAGGCAATGGAATAATACATAACTGAGAAATAGAGAAGTTTACAGATCAAGGTTAGGAACA 660
 QY 661 GAGAGACAGCAAGATATGGGCCAAACAGAGATATCTGTGTAGACAGTTCTCTCCCGCTC 720
 DB 661 GAGAGACAGCAAGATATGGGCCAAACAGAGATATCTGTGTAGACAGTTCTCTCCCGCTC 720
 QY 721 AGGGCCAAAGACAGTTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAGACAG 780
 DB 721 AGGGCCAAAGACAGTTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTAGACAG 780
 QY 781 TTCTGCTCCCGCTCAGGSCCAAGACAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGAC 840
 DB 781 TTCTGCTCCCGCTCAGGSCCAAGACAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGAC 840
 QY 841 GTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGC 900
 DB 841 GTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGTGC 900
 QY 901 CTTATTGTAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCTGTCTCCCG 960
 DB 901 CTTATTGTAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTCCGCGCTTCTGTCTCCCG 960
 QY 961 AGCTCAATAAAGAGCCCAACACCTCTCACTCGGCGGCCAGTCTCTCCGATAGACTGCGT 1020
 DB 961 AGCTCAATAAAGAGCCCAACACCTCTCACTCGGCGGCCAGTCTCTCCGATAGACTGCGT 1020
 QY 1021 CGCCGGGTACCGGTATTTCCCAATAAAGCTTCTGCTGTTTGCATCCGATCGTGACCTC 1080
 DB 1021 CGCCGGGTACCGGTATTTCCCAATAAAGCTTCTGCTGTTTGCATCCGATCGTGACCTC 1080
 QY 1081 GCTGATCCTTTGGAGGGTCTCTCTCAGATTGATTGATCTGCTGCGGGCTTCTTCATT 1140
 DB 1081 GCTGATCCTTTGGAGGGTCTCTCTCAGATTGATTGATCTGCTGCGGGCTTCTTCATT 1140
 QY 1141 TGGAGGTTTCCACCGAGATTGGAGACCCCTGCGCCAGGAGCCACCGACCCCGCGGGA 1200
 DB 1141 TGGAGGTTTCCACCGAGATTGGAGACCCCTGCGCCAGGAGCCACCGACCCCGCGGGA 1200
 QY 1201 GGTAGCTGGCCAGCAACTTATCTGTGCTGTCCGATGCTGTAGTGTCTATGACTGATTT 1260
 DB 1201 GGTAGCTGGCCAGCAACTTATCTGTGCTGTCCGATGCTGTAGTGTCTATGACTGATTT 1260
 QY 1261 TATGCGCTGCGTCTGCTACTAGTAACTAGTCTCTATCTGCGGACCCGCTGGTGG 1320
 DB 1261 TATGCGCTGCGTCTGCTACTAGTAACTAGTCTCTATCTGCGGACCCGCTGGTGG 1320
 QY 1321 AACTGACGAGTTTCGGAACACCCCGCGCAACCTCGGAGACGTCCTCCAGGAGCTTCGGGG 1380
 DB 1321 AACTGACGAGTTTCGGAACACCCCGCGCAACCTCGGAGACGTCCTCCAGGAGCTTCGGGG 1380
 QY 1381 CGGTTTTTGTGCGCCGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTGGTGCAC 1440
 DB 1381 CGGTTTTTGTGCGCCGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTGGTGCAC 1440
 QY 1441 CCCCTTATAGAGGAGGATATGTTCTGCTAGGAGACAGAACCTTAAACAGATTCCCGC 1500
 DB 1441 CCCCTTATAGAGGAGGATATGTTCTGCTAGGAGACAGAACCTTAAACAGATTCCCGC 1500
 QY 1501 CTCCTGTGAATTTTGTCTTTCGGTTTGGGACCCGAAGCCCGCGCTTGTGTCTGCT 1560
 DB 1501 CTCCTGTGAATTTTGTCTTTCGGTTTGGGACCCGAAGCCCGCGCTTGTGTCTGCT 1560
 QY 1561 GAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAATATG 1620
 DB 1561 GAGCATCGTCTGTGTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAATATG 1620
 QY 1621 GCCCGGGCTAGACTGTACCTCTCCCTTAAAGTTTGAACCTTAGCTCACTCGAAAGATGTCG 1680
 DB 1621 GCCCGGGCTAGACTGTACCTCTCCCTTAAAGTTTGAACCTTAGCTCACTCGAAAGATGTCG 1680
 QY 1681 AGCGGATCGCTCACCAACAGTCGGTAGATGTCAAGAGAGACGTTGGGTTTACCTTCTGCT 1740

Db	1681	ACGGATCGCTCA	CAA	CAAGTCGGTAGATGTCA	AAGAAGACGCTTG	TGGGTACCTTTCGT	1741
Qy	1741	CTGCAGAAATGCC	CAACCTTTAA	CGTCGCGATGCGCCG	CAGACGGCACTTTAA	ACCGAGACC	1800
Db	1741	CTGCAGAAATGCC	CAACCTTTAA	CGTCGCGATGCGCCG	CAGACGGCACTTTAA	ACCGAGACC	1800
Qy	1801	TCATCACCCAGGTT	AAGATCAAGTCTTTT	CAACCTGCGCCGCGAT	TGGACATCCAGACCGG	1860	
Db	1801	TCATCACCCAGGTT	AAGATCAAGTCTTTT	CAACCTGCGCCGCGAT	TGGACATCCAGACCGG	1860	
Qy	1861	TCCCTCATCATCGT	CAGCTGGGAAGCTT	TGGCTTTTGAACCC	CCCTCCCTGGGTCAAGCCCT	1920	
Db	1861	TCCCTCATCATCGT	CAGCTGGGAAGCTT	TGGCTTTTGAACCC	CCCTCCCTGGGTCAAGCCCT	1920	
Qy	1921	TTGTACACCCATA	AGCCTCCGCTCTCT	CTCTCCATCGCCCGCT	CTCTCCCTTTGAAC	1980	
Db	1921	TTGTACACCCATA	AGCCTCCGCTCTCT	CTCTCCATCGCCCGCT	CTCTCCCTTTGAAC	1980	
Qy	1981	CTCCTCGTTGCA	CCCGCTCGATCTCC	CTCTTATCCAGCCCT	CACTCTCTTAGGG	2040	
Db	1981	CTCCTCGTTGCA	CCCGCTCGATCTCC	CTCTTATCCAGCCCT	CACTCTCTTAGGG	2040	
Qy	2041	CCCCATATGSC	CAATAGATCTTAT	TGGGCAACCCCG	CCCTGTAACTTCCCTG	2100	
Db	2041	CCCCATATGSC	CAATAGATCTTAT	TGGGCAACCCCG	CCCTGTAACTTCCCTG	2100	
Qy	2101	ACCTTGACATCA	CAAGAGTTACT	TAAAGCCCTCTCT	CAAGCTCACTTACAGCTTCTA	2160	
Db	2101	ACCTTGACATCA	CAAGAGTTACT	TAAAGCCCTCTCT	CAAGCTCACTTACAGCTTCTA	2160	
Qy	2161	CTTAGTCCAGCA	CGAAGTCTGGAGA	CTCTGGCGGACGCT	TACAAGAACTTGGACG	2220	
Db	2161	CTTAGTCCAGCA	CGAAGTCTGGAGA	CTCTGGCGGACGCT	TACAAGAACTTGGACG	2220	
Qy	2221	ACCGGTGGTAC	CTCACCCTTAC	CGAGTCGGCGACA	CAGTGTGGTCCGCGACACGAC	2280	
Db	2221	ACCGGTGGTAC	CTCACCCTTAC	CGAGTCGGCGACA	CAGTGTGGTCCGCGACACGAC	2280	
Qy	2281	TAAGAACTTAGAA	CTCGCTGGAAAGCA	CTTTACA	CAGTCTCTGCTGACCA	CCCGCACCGC	2340
Db	2281	TAAGAACTTAGAA	CTCGCTGGAAAGCA	CTTTACA	CAGTCTCTGCTGACCA	CCCGCACCGC	2340
Qy	2341	CCTCAAAAGTAGA	CGGCATCGCAGT	TCGGATA	CACGCGCCCA	CGTGAAGCTCCCGACCC	2400
Db	2341	CCTCAAAAGTAGA	CGGCATCGCAGT	TCGGATA	CACGCGCCCA	CGTGAAGCTCCCGACCC	2400
Qy	2401	CGGGGTGGACCA	CTCTAGACTGCA	CTGGATGGAGCT	GTATCATCTCTTCTTGGTA	2460	
Db	2401	CGGGGTGGACCA	CTCTAGACTGCA	CTGGATGGAGCT	GTATCATCTCTTCTTGGTA	2460	
Qy	2461	GCAACAGCTAC	CAGGTGTCCACT	CCGACATCCAG	CTGACCCAGACGCC	CAAGCTGAGC	2520
Db	2461	GCAACAGCTAC	CAGGTGTCCACT	CCGACATCCAG	CTGACCCAGACGCC	CAAGCTGAGC	2520
Qy	2521	GCCAGCGTGGT	TGACAGAGTGAC	CAATCACCT	TGAAGCCAGT	CAGGATGTGGGTACTTCT	2580
Db	2521	GCCAGCGTGGT	TGACAGAGTGAC	CAATCACCT	TGAAGCCAGT	CAGGATGTGGGTACTTCT	2580
Qy	2581	GTAGCTTGGTAC	AGAGAGCAGGTA	AGGCTCCAA	AGCTGTGTATCTAT	CTGGAATCC	2640
Db	2581	GTAGCTTGGTAC	AGAGAGCAGGTA	AGGCTCCAA	AGCTGTGTATCTAT	CTGGAATCC	2640
Qy	2641	ACCGGACACTGT	GTGGCCAGCAGAT	TACAGCGGTAG	CGGTAGCGGTACCGACTTCA	2700	
Db	2641	ACCGGACACTGT	GTGGCCAGCAGAT	TACAGCGGTAG	CGGTAGCGGTACCGACTTCA	2700	
Qy	2701	TTTACCATCAG	CAGCCTCCAG	CCAGAGGACAT	CGCCACTACT	CTACTGCCAGCAATATAGC	2760
Db	2701	TTTACCATCAG	CAGCCTCCAG	CCAGAGGACAT	CGCCACTACT	CTACTGCCAGCAATATAGC	2760
Qy	2761	CTCTATCGGTCT	TCGCGCAAGGCA	CAAGTGGAAAT	TCAAA	CGAGGTGCTCAGGATCG	2820
Db	2761	CTCTATCGGTCT	TCGCGCAAGGCA	CAAGTGGAAAT	TCAAA	CGAGGTGCTCAGGATCG	2820

QY	3901	GTCTCCAGAAAAAGGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAG	3960	Db	4981	ATCTCTCTACGCCGAGCATCGTGGCCGGGTACCGAGCTCGAATTCGTAATCATGGTCAT	5040
Db	3901	GTCTCCAGAAAAAGGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAG	3960	QY	5041	AGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATTTCCACAAATACGAGCGGAA	5100
QY	3961	TAAAGCCATTTTGCAAGGCATGGAATAATCATAACTGAGAAATAGAGAAAGTTCCAGATCAAG	4020	Db	5041	AGCTGTTTCTGTGTGAAATTTGTTATCCGCTCAAAATTTCCACAAATACGAGCGGAA	5100
Db	3961	TAAAGCCATTTTGCAAGGCATGGAATAATCATAACTGAGAAATAGAGAAAGTTCCAGATCAAG	4020	QY	5101	GCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACTCAATTAATTTGGTTGC	5160
QY	4021	GTATGGAACAGAGAGACAGCAATATGGGCCAAACAGAGATATCTGTGTAGAGCAGTTCC	4080	Db	5101	GCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACTCAATTAATTTGGTTGC	5160
Db	4021	GTATGGAACAGAGAGACAGCAATATGGGCCAAACAGAGATATCTGTGTAGAGCAGTTCC	4080	QY	5161	GCTCACTGCCGCTTTTCCAGTCGGGAAACCTGTCTGTGCGAGCTGATTAAATGAATCGGCC	5220
QY	4081	TGCCCCGCTCAGGGCCAAAGACAGTTGGAAACAGGAAATATGGGCCAAACAGAGATATCTG	4140	Db	5161	GCTCACTGCCGCTTTTCCAGTCGGGAAACCTGTCTGTGCGAGCTGATTAAATGAATCGGCC	5220
Db	4081	TGCCCCGCTCAGGGCCAAAGACAGTTGGAAACAGGAAATATGGGCCAAACAGAGATATCTG	4140	QY	5221	AAAGCGGGGAGAGCGGTTTGGCGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACT	5280
QY	4141	TGTTAAGCAGTTTCTGTGCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGCGGTCC	4200	Db	5221	AAAGCGGGGAGAGCGGTTTGGCGTATTTGGCGCTCTTCCGCTTCTCGCTCACTGACT	5280
Db	4141	TGTTAAGCAGTTTCTGTGCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGCGGTCC	4200	QY	5281	CGCTCGCTCGTCTCGCTTCCGCTGCGGCGAGGGTATCAGCTCACTCAAGGGCGTAATAC	5340
QY	4201	GCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGCCTGAAT	4260	Db	5281	CGCTCGCTCGTCTCGCTTCCGCTGCGGCGAGGGTATCAGCTCACTCAAGGGCGTAATAC	5340
Db	4201	GCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGAGCCTGAAT	4260	QY	5341	GGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA	5400
QY	4261	GACCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGCTTCTGCTCGCGCTT	4320	Db	5341	GGTTATCCACAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA	5400
Db	4261	GACCTGTGCTTATTTGAACTAAACCAATCAGTTTCTGCTTCTGCTTCTGCTCGCGCTT	4320	QY	5401	AGGCAGAGAAACCGTAAAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCGCCCCCTG	5460
QY	4321	CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGA	4380	Db	5401	AGGCAGAGAAACCGTAAAAAGGCGGCTTGTGCGGCTTTTCCATAGGCTCGCCCCCTG	5460
Db	4321	CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGA	4380	QY	5461	ACGACATCACAAAATCGAGCTCAAGTCAAGGTGCGAAACCCGACAGGACTATAAA	5520
QY	4381	TAGACTGCTGCGCGGGTACCGGTGTTCTCAATAAACCCCTCTTGCAGTTGATCCGACT	4440	Db	5461	ACGACATCACAAAATCGAGCTCAAGTCAAGGTGCGAAACCCGACAGGACTATAAA	5520
Db	4381	TAGACTGCTGCGCGGGTACCGGTGTTCTCAATAAACCCCTCTTGCAGTTGATCCGACT	4440	QY	5521	GATACAGGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGC	5580
QY	4441	CGTGGTCTCGCTTCTTGGGAGGCTCTCTGAGTGATGACTACCCGTCAGCGGGT	4500	Db	5521	GATACAGGCGGTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGC	5580
Db	4441	CGTGGTCTCGCTTCTTGGGAGGCTCTCTGAGTGATGACTACCCGTCAGCGGGT	4500	QY	5581	TTACCGGATACCTGTCCGCTTTTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAC	5640
QY	4501	CTTTCAAGTTTCTCCCACTACACAGGTCTCACTAAACATTCCTGATGTCGCGAGGACTC	4560	Db	5581	TTACCGGATACCTGTCCGCTTTTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAC	5640
Db	4501	CTTTCAAGTTTCTCCCACTACACAGGTCTCACTAAACATTCCTGATGTCGCGAGGACTC	4560	QY	5641	GCCTAGGATCTCTAGTTTCCGTTAGGTGCTTCCGCTCAAGTGGCTGTGTGTCAGCAAC	5700
QY	4561	CGTCAGCCGGTTTTTGTATTAATAAATGCAAGACAGTGTCCCTTCAAGCCAGACT	4620	Db	5641	GCCTAGGATCTCTAGTTTCCGTTAGGTGCTTCCGCTCAAGTGGCTGTGTGTCAGCAAC	5700
Db	4561	CGTCAGCCGGTTTTTGTATTAATAAATGCAAGACAGTGTCCCTTCAAGCCAGACT	4620	QY	5701	CCCCGCTTCCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTCTTGAGTCCAAACCCGG	5760
QY	4621	ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACTATCTGCCACA	4680	Db	5701	CCCCGCTTCCAGCCGACCGCTGCGCTTATCCGTTAACTATCGTCTTGAGTCCAAACCCGG	5760
Db	4621	ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACTATCTGCCACA	4680	QY	5761	TAAAGACAGACTTATCGGCACTGGGACAGCCACTGGTAAACAGGATTAGCAGCGAGGT	5820
QY	4681	CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTCGTCTA	4740	Db	5761	TAAAGACAGACTTATCGGCACTGGGACAGCCACTGGTAAACAGGATTAGCAGCGAGGT	5820
Db	4681	CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCCTCGTCTA	4740	QY	5821	ATGAGGCGGTCTACAGAGTTTGAAGTGTGCGCTAACTACGCTACACTAGAAGA	5880
QY	4741	TTTTGCTGCTTCTCGATTTTAAATTTCTAGTTTGAATTTAGTCAATCTATCCTCGTCTA	4800	Db	5821	ATGAGGCGGTCTACAGAGTTTGAAGTGTGCGCTAACTACGCTACACTAGAAGA	5880
Db	4741	TTTTGCTGCTTCTCGATTTTAAATTTCTAGTTTGAATTTAGTCAATCTATCCTCGTCTA	4800	QY	5881	CAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCCGAAAAAGAGTTGGTAGCT	5940
QY	4801	GATTCAGAGTAGTTTAAATCTGAGGGCAGGCTCTGTGAAAGGTTCCCTGGGCTCAG	4860	Db	5881	CAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCCGAAAAAGAGTTGGTAGCT	5940
Db	4801	GATTCAGAGTAGTTTAAATCTGAGGGCAGGCTCTGTGAAAGGTTCCCTGGGCTCAG	4860	QY	5941	CTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTGGTTGCAAGACGACAGA	6000
QY	4861	TGTGAGATTTGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGGAATCTAGA	4920	Db	5941	CTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTGGTTTTTGGTTGCAAGACGACAGA	6000
Db	4861	TGTGAGATTTGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGGAATCTAGA	4920	QY	6001	TTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTGTGATCTTTTCTACGGGCTCTGACG	6060
QY	4921	GTGCATCTCAGAGTCCCGGGTCCGGGCTCTGATCTCAGGGATCTTTTGGCTAGAG	4980	Db	6001	TTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTGTGATCTTTTCTACGGGCTCTGACG	6060
Db	4921	GTGCATCTCAGAGTCCCGGGTCCGGGCTCTGATCTCAGGGATCTTTTGGCTAGAG	4980	QY	6061	CTCAGTGAAACGAAACCTCAGCTTAAAGGGATTTTGGTTCATGAGATTATCAAAAGGATCT	6120
QY	4981	ATCCTCTACCGGACCGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGGTCAT	5040				

```
Db 6061 CTCAGTGGAAAGAACTCACGTTAAGGGAATTTGGTCATGAGATTATCAAAAGGATCT 6120
Qy 6121 TCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATAAGTATATAGAT 6180
Db 6121 TCACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAATCAATAAGTATATAGAT 6180
Qy 6181 AAACCTTGCTGACAGATTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTC 6240
Db 6181 AAACCTTGCTGACAGATTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTC 6240
Qy 6241 TATTTTCGTTTCATCCATAGTTCCTGACCTCCCGCTCGTGTAGATAACTACGATACGGGAGG 6300
Db 6241 TATTTTCGTTTCATCCATAGTTCCTGACCTCCCGCTCGTGTAGATAACTACGATACGGGAGG 6300
Qy 6301 CGTTACCAATCGGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCAACCGGCTCCAG 6360
Db 6301 CGTTACCAATCGGCCCCAGTGTGCAATGATACCGCGAGACCCAGCTCAACCGGCTCCAG 6360
Qy 6361 ATTTATCAGCAATAAACAGCCAGCCGGAAGGCGGAGCGAGAGTGGTCTCGCACTT 6420
Db 6361 ATTTATCAGCAATAAACAGCCAGCCGGAAGGCGGAGCGAGAGTGGTCTCGCACTT 6420
Qy 6421 TATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAGTAGTTGCGCCAG 6480
Db 6421 TATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTAGAGTAGTAGTTGCGCCAG 6480
Qy 6481 TTAATAGTTTCGCAACGTTGTCGCAATGCTACAGGCTCGTGGTGTCAAGCTCGTCTGTT 6540
Db 6481 TTAATAGTTTCGCAACGTTGTCGCAATGCTACAGGCTCGTGGTGTCAAGCTCGTCTGTT 6540
Qy 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGGGAGTTACATGATCCCCAT 6600
Db 6541 TGGTATGCTTCATTCAGCTCGGTTCCCAAGATCAAGGGGAGTTACATGATCCCCAT 6600
Qy 6601 GTTGTGAAAAAGCGTTAGCTCTTCGCTCGCTCCGATCGTGTGCAAGTAGTAGTTGGC 6660
Db 6601 GTTGTGAAAAAGCGTTAGCTCTTCGCTCGCTCCGATCGTGTGCAAGTAGTAGTTGGC 6660
Qy 6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCAATGCCATC 6720
Db 6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTACTGTCAATGCCATC 6720
Qy 6721 CGTAAAGTCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGAGAAATAGTGTAT 6780
Db 6721 CGTAAAGTCTTTCTGTGACTGTGTAGTACTCAACCAAGTCAATCTGAGAAATAGTGTAT 6780
Qy 6781 GCGGCGACCGAGTTCCTTCGCGGCGTCAATAGGGATAATCCGCGCCACATAGCAG 6840
Db 6781 GCGGCGACCGAGTTCCTTCGCGGCGTCAATAGGGATAATCCGCGCCACATAGCAG 6840
Qy 6841 AACTTTAAAGTGTCTCATCATTTGAAACGTTCTTCGGGGGAAACCTCTCAAGATCTT 6900
Db 6841 AACTTTAAAGTGTCTCATCATTTGAAACGTTCTTCGGGGGAAACCTCTCAAGATCTT 6900
Qy 6901 ACCGCTGTGAGATCCAGTTCGATTAACCACTCGTSCACCCCACTGATCTTCAGCATC 6960
Db 6901 ACCGCTGTGAGATCCAGTTCGATTAACCACTCGTSCACCCCACTGATCTTCAGCATC 6960
Qy 6961 TTTTACTTTTACCAAGCGTTTCGCGTGGAGCAAAACAGGAAGGCAAAATGCGCAAAAAA 7020
Db 6961 TTTTACTTTTACCAAGCGTTTCGCGTGGAGCAAAACAGGAAGGCAAAATGCGCAAAAAA 7020
Qy 7021 GGGAAATAGGGCGACACGGAATGTTGAATACTCATCTCTCTCTTTTCAATATTATTG 7080
Db 7021 GGGAAATAGGGCGACACGGAATGTTGAATACTCATCTCTCTCTTTTCAATATTATTG 7080
Qy 7081 AAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTAGAAAAA 7140
Db 7081 AAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTAGAAAAA 7140
Qy 7141 TAAACAAATAGGGTTTCGCGCACATTTCCCGCAAAAGTGCACCTGACGCTTAAGAAAC 7200
Db 7141 TAAACAAATAGGGTTTCGCGCACATTTCCCGCAAAAGTGCACCTGACGCTTAAGAAAC 7200
```

```
Qy 7201 CATTATTATCATGATCACTTAACCTATAAAATAGCGGTATCAAGAGGCCCTTTCTGCTCGC 7260
Db 7201 CATTATTATCATGATCACTTAACCTATAAAATAGCGGTATCAAGAGGCCCTTTCTGCTCGC 7260
Qy 7261 GCGTTTCGGTGCATCAGCGTGAAACCTCTGACACATGCACTCCCGGAGACGGTCACAGC 7320
Db 7261 GCGTTTCGGTGCATCAGCGTGAAACCTCTGACACATGCACTCCCGGAGACGGTCACAGC 7320
Qy 7321 TTGTCGTGTAACGGATGCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGGTGTGG 7380
Db 7321 TTGTCGTGTAACGGATGCGGGAGCAGACAAGCCCGTCAGGCGCGCTCAGCGGGTGTGG 7380
Qy 7381 CGGTTGTCGGGGCTGGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACA 7440
Db 7381 CGGTTGTCGGGGCTGGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACA 7440
Qy 7441 TATCGGCTGTGAATATCCGCAAGATCGGTAAAGGAGAAATACCGCATCAGGCGCCATTC 7500
Db 7441 TATCGGCTGTGAATATCCGCAAGATCGGTAAAGGAGAAATACCGCATCAGGCGCCATTC 7500
Qy 7501 GCCATTCAAGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTATTACG 7560
Db 7501 GCCATTCAAGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTCTTCGCTATTACG 7560
Qy 7561 CCAGCTGCGGAAAGGGGATGCTGCAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTC 7620
Db 7561 CCAGCTGCGGAAAGGGGATGCTGCAAGCGGATTAAGTTGGGTAAACCGCAGGGTTTC 7620
Qy 7621 CCAGTCAAGAGTGTGTAACAGCAGCGCCAGTGCC 7654
Db 7621 CCAGTCAAGAGTGTGTAACAGCAGCGCCAGTGCC 7654

RESULT 3
ABS56664
ID ABS56664 standard; DNA; 9320 BP.
AC ABS56664;
XX
XX 24-JAN-2003 (first entry)
XX
DE Plasmid p5-Puro-CMV- (N'-EGFP) -CMV-Red (EGFP-EJ) DNA.
XX
XX Vector; genotoxicity; food component; cosmetic; medicine; cancer;
XX genetic stability; hygromycin; ds.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..1592
FT /tag= a
FT /note= "retroviral vector p5NM"
FT CDS 1617..2216
FT /tag= b
FT /product= "puromycin resistance gene from pRetron"
FT promoter 2267..2848
FT /tag= c
FT /note= "CMV promoter from pEGFP-N1"
FT CDS 2906..3348
FT /tag= d
FT /product= "N'-EGFP"
FT misc_feature 3374..3392
FT /tag= e
FT /note= "retroviral vector p5NM"
FT promoter 3411..3992
FT /tag= f
FT /note= "CMV promoter from pDsRed1-N1"
FT CDS 4038..4718
FT /tag= g
FT /product= "pDsRed1-N1"
FT CDS 4766..5508
FT /tag= h
```


Db 9200 GGGTGTGGCGGTGTCGGGGCTGGCTTAACCTATGGGCATCAGAGCAGATTGTACTGAG 9259
 QY 7432 AGTGACCATATGCGGTGTGAATACCGCACAGATCGGTAAAGAGAAATACCGATCAG 7491
 Db 9260 AGTGACCATATGCGGTGTGAATACCGCACAGATCGGTAAAGAGAAATACCGATCAG 9319
 QY 7492 G 7492
 Db 9320 G 9320
 RESULT 4
 AAX33181
 ID AAX33181 standard; DNA; 6644 BP.
 XX
 AC AAX33181;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Base sequence of the plasmid pRx-ires-bsr.
 XX
 KW Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
 KW bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
 KW autoimmune disease; graft rejection reaction; inflammation;
 KW inflammatory disease; ss.
 XX
 OS Synthetic.
 OS Cowpox virus.
 XX
 PN WO9913073-A2.
 XX
 PD 18-MAR-1999.
 XX
 PF 07-SEP-1998; 98WO-JP04010.
 XX
 PR 08-SEP-1997; 97JP-00259235.
 XX
 XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
 XX
 PI Hamada H;
 XX
 DR WPI; 1999-243728/20.
 XX
 XX New apoptosis-resistant virus-sensitive cell.
 PT
 XX
 PS Example 1; Page 38-41; 51pp; English.
 CC
 CC The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis by gene transfer, or where the
 CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC therapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present gene represents the base sequence of the plasmid
 CC pRx-ires-bsr, which contains the cowpox virus bar gene, and is used in an
 CC example from the present invention
 XX
 SQ Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
 Query Match 43.2%; Score 3307.8; DB 2; Length 6644;

Best Local Similarity 72.9%; Pred. No. 0;
 Matches 4739; Conservative 0; Mismatches 1352; Indels 410; Gaps 21;
 QY 843 TTCTAGAGAACCATCAGATGTTTCCAGGGTCCCCCAAGGACCTGAATGACCCCTGTGCT 902
 Db 467 TGCTAGAGAACCATCAGATGTTTCCAGGGTCCCCCAAGGACCTGAATGACCCCTGTGCT 526
 QY 903 TATTTGAACTAACCAATCAGTTCGGTCTCTCGTCTTCTGTTCCGGCGCTTCTGCTCCCGGAG 962
 Db 527 TATTTGAACTAACCAATCAGTTCGGTCTCTCGTCTTCTGTTCCGGCGCTTCTGCTCCCGGAG 586
 QY 963 CTCATATAAAGAGAGCCCAACCCCTCACTCGGGCGCGCAGTCTCTCCGATAGACTCGGTG 1022
 Db 587 CTCATATAAAGAGAGCCCAACCCCTCACTCGGGCGCGCAGTCTCTCCGATAGACTCGGTG 646
 QY 1023 CCGGGGTACCGGTATTCCTCAATAAAGCCTCTTGCTGTTTGGCATCGAATCGGACTCGC 1082
 Db 647 CCGGGGTACCGGTATTCCTCAATAAAGCCTCTTGCGAG-TTGGCATCGACTTGGGTCTGCG 705
 QY 1083 TGATCTTTGGGAGGGTCTCTCAGATTGATTGACTGCCC-ACCTCGGGGGTCTTTCAATTT 1141
 Db 706 TGTTCCTTTGGGAGGGTCTCTCTGAGTGAATTGACTACCGTCAGCGGGGTCTTTCAATTT 765
 QY 1142 GGAGGTTCCACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAG 1201
 Db 766 GGGGGCTCGTCCGGGATCGGGAGACCCCTGCCAGGACCAACCGA-CCACCCACCGGGAG 824
 QY 1202 GTAAGCTGGCCAGCAACTTATCTGTCTGTGCTGCTGCTAGTCTATGACTGATTTT 1261
 Db 825 GTAAGCTGGCCAGCAACTTATCTGTCTGTGCTGCTGCTAGTCTATGACTGATTTT 884
 QY 1262 ATGCGCCTGCGTCCGTACTAGTTAGCTAACTAGTCTGTATCTGCGGACCCGTGTGGA 1321
 Db 885 ATGCGCCTGCGTCCGTACTAGTTAGCTAACTAGTCTGTATCTGCGGACCCGTGTGGA 944
 QY 1322 ACTGACGAGTTTGGAAACACCCGGCCGCAACCCCTGGGAGACGTCCTCAGGAGACTTCGGGGGC 1381
 Db 945 ACTGACGAGTTTGGAAACACCCGGCCGCAACCCCTGGGAGACGTCCTCAGGAGACTTCGGGGGC 1004
 QY 1382 CGTTTTTGTGGCCGACCTGAGTCTTAATAATCCCGATCGTTTAGACTCTTTGGTGCACC 1441
 Db 1005 CGTTTTTGTGGCCGACCTGAGTCCAAAATCCCGATCGTTTAGACTCTTTGGTGCACC 1064
 QY 1442 CCCCTTAGAGAGGAGATATGTGTTCTGTAGGAGACGAGAACCTAAACAGTTCGCCGC 1501
 Db 1065 CCCCTTAGAGAGGAGATATGTGTTCTGTAGGAGACGAGAACCTAAACAGTTCGCCGC 1124
 QY 1502 TCCGTCTGAATTTTGTCTTTCGGTTTGGACCGAAGCCGCGCGCTTTGTCTGCTG 1561
 Db 1125 TCCGTCTGAATTTTGTCTTTCGGTTTGGACCGAAGCCGCGCGCTTTGTCTGCTG 1184
 QY 1562 GAGCATGTTCTGTGTTGTCTCTGTCTGATGTTGTTCTGTATTTGTCTGAAAATATGGG 1621
 Db 1185 GAGCATGTTCTGTGTTGTCTCTGTCTGATGTTGTTCTGTATTTGTCTGAAAATATGGG 1244
 QY 1622 CCGGGCTAGACTGTTTACACCTCCCTTAAGTTTGAACCTTAGTCTACTGGAAGATGTGCA 1681
 Db 1245 CC-----AGACTGTTTACACCTCCCTTAAGTTTGAACCTTAGTCTACTGGAAGATGTGCA 1298
 QY 1682 GCGGATCGCTCAACACCATCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGCTC 1741
 Db 1299 GCGGATCGCTCAACACCATCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGCTC 1358
 QY 1742 TGCAGATGGCCAACTTTTAACGTCGGATGGCGGAGACGCGACCTTTAACCGAGACCT 1801
 Db 1359 TGCAGATGGCCAACTTTTAACGTCGGATGGCGGAGACGCGACCTTTTAACCGAGACCT 1418
 QY 1802 CATCACCCAGGTTAAGATCAAGTCTTTTACCTCGCCCGCATGACACCCAGACAGGT 1861
 Db 1419 CATCACCCAGGTTAAGATCAAGTCTTTTACCTCGCCCGCATGACACCCAGACAGGT 1478
 QY 1862 CCCCTTACATCGTGACCTGGGAGCCCTTGGCTTTTGAACCCCTCCCTCGGCTCAAGCCCTT 1921

Db 1479 CCCCTACATCGTACCTGGGAAGCCCTTGGCTTTTGACCCCTCCTCGGTCAAGCCCTT 1538
Qy 1922 TGTACACCTTAAGCCTCGCCTCTCTCTTCCTCAATCCGCCCGCTCTCCTCCCTTGAACC 1981
Db 1539 TGTACACCTTAAGCCTCGCCTCTCTCTTCCTCAATCCGCCCGCTCTCCTCCCTTGAACC 1598
Qy 1982 TCCTCGTTTCGACCCCGCTCGATCTCTCTCTTTATCCAGCCTCACTCTCTCTCTAGGCGC 2041
Db 1599 TCCTCGTTTCGACCCCGCTCGATCTCTCTCTTTATCCAGCCTCACTCTCTCTCTAGGCGC 1658
Qy 2042 CCCCATATGGCAATAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101
Db 1659 -----CCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 1708
Qy 2102 CCCTGACATGACAAGAGTTATACAGACCCCTCTCTCCAAGCTCACTTACAGGCT-TCTA 2160
Db 1709 CCCTGACATGACAAGAGTTATACAGACCCCTCTCTCCAAGCTCACTTACAGGCTCTCTA 1768
Qy 2161 CTTAGTCAGACGAGTCTGGAGACCTCTGGCGGACCTTACCAAGAACCACTGGACCG 2220
Db 1769 CTTAGTCAGACGAGTCTGGAGACCTCTGGCGGACCTTACCAAGAACCACTGGACCG 1828
Qy 2221 ACCGGTGTACTCACTCACTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACCGAG 2280
Db 1829 ACCGGTGTACTCACTCACTTACCGAGTCGGGACACAGTGTGGTCCGCCGACACCGAG 1888
Qy 2281 TAAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGACCAACCCCAACCGC 2340
Db 1889 TAAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGACCAACCCCAACCGC 1948
Qy 2341 CCTCAAAGTAGACGGCATCGAGCTTGGATACACGCCGCCACAGTGAAGGCTGGCGACCC 2400
Db 1949 CCTCAAAGTAGACGGCATCGAGCTTGGATACACGCCGCCACAGTGAAGGCTGGCGACCC 2008
Qy 2401 CGGGGTGGACCATCTCTAGACTGCCATGGATGGAGTCTATCATCTCTCTTCTGTGTA 2460
Db 2009 CGGGGTGGACCATCTCTAGACTGCCAGTCCAGTGTGGTGGTAGGGAATTCAGCTC 2068
Qy 2461 GCAACAGCTACAGTGTCTCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTCAGC 2520
Db 2069 GGTACCCGGGGGCGCGCGATCTCTTAATTAAGCGCGCGGTAAACA----- 2116
Qy 2521 GCCAGCTGGGTGACAGAGTGACCATACCTGTGAAGGCCAGTCAGGATGGGTACTTCT 2580
Db 2117 -----TTG 2119
Qy 2581 GTAGCTTGGTACACAGACAGGTAAGGCTCCAAAGCTGCTGATCTACTGACATCC 2640
Db 2120 TTAATACTAACTAAGTAGACAGGTTTCCCTCTAGCGGGATCAATTCCTCCCTCTCCC 2179
Qy 2641 ACCCGGCACACTGTGTGCCAAGCAGATTCAGCGGTAGCGGTACCGACTTCACC 2700
Db 2180 TCCCTCCCTTAACTTACTGGCGA----- 2206
Qy 2701 TTCACCATCAGAGCTCCAGCCAGAGACATCGCACCTACTACTCGCAGCAATATAGC 2760
Db 2207 -----AGCCGCTTGGAAATAAGCCGGTGTGCTTGTCTATATGTTATTTTCCACC 2257
Qy 2761 CTCTATCGTCTCGGCCAGGACCAAGGTGGNAATCAACGAGGTGGCTCAGGATCG 2820
Db 2258 ATATGCGCTCTTTGGCAATGTGA----- 2282
Qy 2821 GGTGGATCCGCTCTGGTGGCTCAGGATCGGAGGTCCAACTGGTGGAGAGCGGTGAGGT 2880
Db 2283 -----GGCCCGGAAACCTG 2297
Qy 2881 GTTGTGCAACTGSCCGCTCCCTGCGCTGCTCTGCTCCCGATCTGGCTTCGATTTCAAC 2940
Db 2298 GCGCTGTCTTCTGACGAGCATCTCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAG 2357
Qy 2941 ACATATTGGATGAGTTGGGTGAGACAGCAGCCTCGAAAAGTCTTGTAGTGGATTTGGAGAA 3000
Db 2358 GTCTGTTGAATGTCGTGAAGGAAGCAGTTCTCTCTGGAGCTTCTTGAAGACAAACAGT 2417

Qy 3001 ATTCAATCCAGATACAGTACGATTAACTATATCGCGCTCTCTAAAGGATAGATTACAAATA 3060
Db 2418 CTGTAGCGA---CCCTTTGCAGGCAGCGGAACCCCACTTGGCGACAGGTGCTCTCGC 2474
Qy 3061 TCGGAGACACGCGCAAGACACATTTGCTCTGCAATGGACAGCTTGAAGCCGAGAAC 3120
Db 2475 GCCAAAAGCCACGTGTATAAGATACACTGCAAAAGGGGCGACACCCCACTGCGACGTTG 2534
Qy 3121 ACCGGGGTCTATTTTGTGCAAGCCTTACTTTCGGCTTCCCTCGTGTGCTTATTTGGGGC 3180
Db 2535 TGAGTTGATAGTTGTGGAAGAGTCAATGGCTCTCTCAAGCGTATTTCAACAGGGGC 2594
Qy 3181 CAAGGGACCCCGGTCAACGCTCTCAGTGTCAAGCCCAACGACGCGAGCGCGGACCA 3240
Db 2595 TGAAGGATGCCAGAAAGGTACCCCAAT----- 2621
Qy 3241 CCAACACGGGCGCCCAACCATCGCTCGCAGCCCTGTCTGCGCCAGAGGGCGCTCGG 3300
Db 2622 -----GTATGG 2628
Qy 3301 CCAGCGCGGGGCGCGAGTGCACACGAGGGGGCTGGACTTTCGCCCTGGATCCCAAACTC 3360
Db 2629 ATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGTTAAAAAGCT 2688
Qy 3361 TGCTACCTGTGGATGGAATCTCTCTCATCTATGCTGTCTCTCTCTCTCTCTCTCTCTG 3420
Db 2689 CTAGGCCCCCGAA-----CCACGGGAGCGTGTCTCTCTTTGMAAAAACACGATAATA 2741
Qy 3421 AGAGTGAAGTTACAGAGGCGGAGAGCGCCCGCTTACAGAGGGGCGGACCAACAGCTC 3480
Db 2742 CCATGGTTCATGAAAACATTTAAACATTTCTCAACAGATCTAGAATTAGTAAGTAGCGA 2801
Qy 3481 TATAACGAGTCAATCTAGGACGAAGAGAGGAGTACGATGTTTTGGACAAGAGCGTGGC 3540
Db 2802 CAGAGAAGTTACATCTCTTATGAGGATATAAACAATCATCTGCGAGCGCAATTCGTA 2861
Qy 3541 CGGACCTCGAGATGGGGGAAAGCCGAGAGGAAGAACCTTCAGGAGGCTCTGTACAAT 3600
Db 2862 CGAAAACAGGAGAAATCATTTTCGCGAGTACATATTTGAAGCGTATATAGGACGAGTAACTG 2921
Qy 3601 GAATCGCAGAAAGATAGATGGCGGAGCGCTTACAGTGAAGTTGGGATGAAGGCGAGCGC 3660
Db 2922 TTTGTGCAAGACCATTTGCGATTGGTAG--TGCAGTTTCGAATGGACAAAAGGATTTTGA 2979
Qy 3661 CGGAGGGCAAGGGCGCAGATGGCTTTTACAGAGTCTCAGTACAGCCACCAAGGACACC 3720
Db 2980 CACGATTTAGCTGTTAGACACCTTATTTGAGCAAGTAGATAGAGTATTCAGTGGT 3039
Qy 3721 TACGACGCCCTTCACATGCGAGGCCCTGCCCCCTCGCTAACTCGACCGCGCGCGATCCG 3780
Db 3040 AAGTCTTTGGTATGTAGG-----GAGTTGATTTTCAGACTATGCACCA 3085
Qy 3781 GATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCTCAGGCTCTAGTTTGTACTAAC 3840
Db 3086 GATTGTTTGTGTTAATAGAAAATGAATGCGAAGTAGTCTCAAAACTACGATTTGAAGAACCTC 3145
Qy 3841 AATATCACCAGCTGAACCTTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTTA 3900
Db 3146 ATTCCACTCAATATATACC--GAATTAAGTTCGACGATATAAATAAAGATTTTATTTA 3203
Qy 3901 GTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960
Db 3204 GTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3263
Qy 3961 TAAGCCCATTTTGAAGGCTAGG--AAATACATAACTGAGATAAGAGATTCAGATCAA 4019
Db 3264 TAAGCCCATTTTGAAGGCTAGGAAAAATATATAACTGAGAAATAGAGAAATTCAGATCAA 3323
Qy 4020 GGTTAGGAAACAGA-GAGACAGCAGAAATATGGCCCAACAGATATCTGTGTAGCAGT 4078
Db 3324 GGTCAAGAAACAGATGGAAACAGCTGAATATGGGGCAACAGGATATCTGTGTAGCAGT 3383

5544 AATCAATCTAAAGTATATAGTAAACTTGGTCTGACAGTTTACCAATGCTTAAATCAGTG 5603
6217 AGGCACCTATCTCAGCGATCTGTCTATTTGGTTATCCATAGTTCGCTGACCTCCCGTCG 6276
5604 AGGCACCTATCTCAGCGATCTGTCTATTTGGTTATCCATAGTTCGCTGACCTCCCGTCG 5663
6277 TGTAGATTAACATACGATACGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGC 6336
5664 TGTAGATTAACATACGATACGGAGGCTTACCATCTGGCCCAAGTCTGCAATGATACCGC 5723
6337 GAGACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGC 6396
5724 GAGACCCACGCTCACCGCTCCAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGC 5783
6397 AGCGCAGAGTGTCTCGCACTTTATCGCCCTCCATCAGTCTATTAATTTGTTCCGGG 6456
5784 AGCGCAGCGGTCTCGCACTTTATCGCCCTCCATCAGTCTATTAATTTGTTCCGGG 5843
6457 AAGCTAGAGTAGTGTGCGCAGTAAATAGTTTGGCAAGTGTGCTGCTGCTACAG 6516
5844 AAGCTAGAGTAGTGTGCGCAGTAAATAGTTTGGCAAGTGTGCTGCTGCTACAG 5903
6517 GC-TCGTGTGTACGCTCGTGTGTGTATGGCTTCAATCAGCTCCGGTCCCAACGAT 6575
5904 GCATCGTGTGTACGCTCGTGTGTGTATGGCTTCAATCAGCTCCGGTCCCAACGAT 5963
6576 CAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCTTCGTCCTC 6635
5964 CAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGTTAGCTCTTCGTCCTC 6023
6636 CGATCGTGTGCAAGTAGTGTGCGCGCATGTTATCACTCATGCTTATGCGACGACTGC 6695
6024 CGATCGTGTGCAAGTAGTGTGCGCGCATGTTATCACTCATGCTTATGCGACGACTGC 6083
6696 ATAATCTCTTACTCTCATGCTCATCGTAAGATGCTTTTCTGTGACTGCTGACTCA 6755
6084 ATAATCTCTTACTCTCATGCTCATCGTAAGATGCTTTTCTGTGACTGCTGACTCA 6143
6756 CCAAGTCACTCTGAAATAGTGTATGCGCGCAGCGAGTTGCTTCTTGGCCGCGTCAATAC 6815
6144 CCAAGTCACTCTGAAATAGTGTATGCGCGCAGCGAGTTGCTTCTTGGCCGCGTCAATAC 6203
6816 GGGATAATACCGCGCCACATGACGAACTTTAAAGTCTCATCATTTGGAAAAAGCTTCTT 6875
6204 GGGATAATACCGCGCCACATGACGAACTTTAAAGTCTCATCATTTGGAAAAAGCTTCTT 6263
6876 CGGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC 6935
6264 CGGCGCAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC 6323
6936 GTGCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAA 6995
6324 GTGCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAA 6383
6996 CAGGAAGCAAAATGCCCAAAAAGGAATAAGGCGCAGCAATGTTGTAATCTCA 7055
6384 CAGGAAGCAAAATGCCCAAAAAGGAATAAGGCGCAGCAATGTTGTAATCTCA 6443
7056 TACTCTCTCTTTTCAATATTTATGAACATTTATCAGGTTATGCTCATGAGCGGAT 7115
6444 TACTCTCTCTTTTCAATATTTATGAACATTTATCAGGTTATGCTCATGAGCGGAT 6503
7116 ACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTTCCGCGCATCTTCCCGAA 7175
6504 ACATATTTGAATGATTTAGAAAAATAAACAATAGGGTTTCCGCGCATCTTCCCGAA 6563
7176 AAGTGCCACCTGACGCTCTAAGAAACCATTTATTCATGACATTAACCTTATTAATAATAGG 7235
6564 AAGTGCCACCTGACGCTCTAAGAAACCATTTATTCATGACATTAACCTTATTAATAATAGG 6623
7236 GTATCAGAGGCGCTTTCGTC 7256
6624 GTATCAGAGGCGCTTTCGTC 6644

RESULT 5

ADAL12886
ID ADAL12886 standard; DNA; 6505 BP.
XX
AC ADAL12886;
XX
DT 20-NOV-2003 (first entry)
XX Murine MSCV retrovirus vector sequence.
DE
DE
KW ds; virus vector; insulator; gypsy; cHS4 dimer core sequence;
KW viral regulatory control element; HScI; stem cell; antianaemic;
KW antiskilling; cytosolic; antiparkinsonian; neurotropic; neuroprotective;
KW gene therapy; thalassaemia; sickle cell anaemia; leukaemia; malignancy;
KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
KW multiple sclerosis; mouse.
XX
OS Mus sp.
XX
FN WO2003070958-A1.
XX
PD 28-AUG-2003.
XX
PF 19-FEB-2003; 2003WO-CA000229.
XX
PR 19-FEB-2002; 2002CA-02370841.
PR 23-FEB-2002; 2002US-0358933P.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Yao S, Ellis J;
XX
XX WPI; 2003-663855/62.
XX
XX New nucleic acid molecule for a virus vector with silencer-blocking
insulator activity, for stem cell marking, stem cell genetic manipulation
studies, gene therapy, or treating thalassemia, sickle cell anemia or
leukemia.
XX
XX Example 1; Page 51-55; 64pp; English.
XX
XX The invention relates to a novel nucleic acid for a virus vector with
silencer-blocking insulator activity that integrates into the genome of a
mammalian stem cell comprising an insulator element, a viral regulatory
control element, and a coding nucleic acid molecule operatively
associated with the viral regulatory element and capable of expression in
the cell. The insulator element comprises a gypsy sequence, a cHS4 dimer
core sequence, or all or part of a fully defined sequence of 611 or 722
base pairs given in the specification. The viral regulatory control
element comprises an infectious retrovirus vector sequence or a
lentivirus vector sequence. The infectious retrovirus sequence comprises
HScI with a fully defined sequence of 6143 bp, given in the
specification. The coding nucleic acid molecule is a reporter gene, where
the reporter gene is PGK eGFP. The host cell is a stem cell of embryonic
or adult tissue origin. The nucleic acid of the invention has
antianaemic, antiskilling, cytosolic, antiparkinsonian, neurotropic, and
neuroprotective activity. The vector may have a use in gene therapy. The
nucleic acid is useful for stem cell marking or stem cell genetic
manipulation studies. The nucleic acid or vectors and compositions
comprising it are also useful in stem cell gene therapy, or for treating
patients with diseases, disorders or abnormal physical states of stem
cells, e.g. thalassaemia, sickle cell anaemia, leukaemia, malignancies,
or neurodegenerative diseases such as Parkinson's disease, Alzheimer's
disease or multiple sclerosis. Host cells comprising the nucleic acid are
useful as research tools to measure levels of expression of the coding
nucleic acid molecule and the activity of the polypeptide encoded by the
coding nucleic acid molecule. The present sequence represents the murine
MSCV retrovirus vector sequence.
XX
XX Sequence 6505 BP; 1382 A; 1732 C; 1666 G; 1464 T; 0 U; 261 Other;

Query Match	42.4%	Score 3244.4	DB 9	Length 6505	
Best Local Similarity	71.1%	Pred. No. 0			
Matches 5064	Conservative	0	Mismatches 1020	Indels 1035	Gaps 25
Qy	552	TGAAGACCCACCTGTAGCTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCGAAGGCAT	611		
Db	1	TGAAGACCCACCTGTAGCTTTGGCAAGCTAGCTTAAGTAAGCCATTTTTCGAAGGCAT	60		
Qy	612	GGAAATACATAACTAGAAATAGAGAAGTTTCAGATCAAGGTTAGGAACACAGAGACAGCA	671		
Db	61	GGAAATACATAACTAGAAATAGAGAAGTTTCAGATCAAGGTTAGGAACACAGAGACAGC	119		
Qy	672	GAATATGGGCCAACACAGGATATCTGTGGTAAGCAGTTCTGCCCGCTCAGGGCCAAAGAA	731		
Db	120	-----	119		
Qy	732	CAGTTGGAACAGGAGATATGGGCCAACACAGGATATCTGTGGTAAGCAGTTCTGCCCG	791		
Db	120	-----AGAAATATGGGCCAACACAGGATATCTGTGGTAAGCAGTTCTGCCCG	166		
Qy	792	GCTCAGGGCCAAAGAACAGATGGTCCCAGATGGGTCCCAGCTTCAGCAGTTCTTAGAGA	851		
Db	167	GCTCAGGGCCAAAGAACAGATGGTCCCAGATGGGTCCCAGCTTCAGCAGTTCTTAGAGA	226		
Qy	852	ACCATCAGATGTTTCCAGGGTGCCCCAAGACCTTG-AAATGACCTGTGCTTATTTGAA	910		
Db	227	ACCATCAGATGTTTCCAGGGTGCCCCAAGACCTTGAAATGACCTGTGCTTATTTGAA	286		
Qy	911	CTAACCAATCAGTTTCGCTTCTCGTTCGTGTTTCGGCGCTTCTGCTCCCCAGCTCAATAA	970		
Db	287	CTAACCAATCAGTTTCGCTTCTCGTTCGTGTTTCGGCGCTTCTGCTCCCCAGCTCAATAA	346		
Qy	971	AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTTCGTCGCCCGGGA	1030		
Db	347	AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTTCGTCGCCCGGGA	406		
Qy	1031	CCGTAATTCCTCAATAAAGCCTTTGCTGTTTGATCCGAATCGTGGACTCGCTGATCCTT	1090		
Db	407	CCGTAATTCCTCAATAAAGCCTTTGCTGTTTGATCCGAATCGTGGACTCGCTGATCCTT	466		
Qy	1091	GGAGGGTCTCTCAGATTAATGACTTGCCCACTCGGGGTCTTTCATTTGAGGTTCC	1150		
Db	467	GGAGGGTCTCTCAGATTAATGACTTGCCCACTCGGGGTCTTTCATTTGAGGTTCC	526		
Qy	1151	ACCGAGATTTGGAGACCCCTGCCAGGGAACACGACCCCGCCGGGAGGTAAGCTGG	1210		
Db	527	ACCGAGATTTGGAGACCCCTGCCAGGGAACACGACCCCGCCGGGAGGTAAGCTGG	586		
Qy	1211	CCAGCAACTTATCTGTGCTGTGTC-----CGATGTCGTAGTGTCTATGACTGATTTTA	1262		
Db	587	CCAGCGTCTGTTGCTGCTGCTCTGCTCTTTGTCGTGTTGTCGGCATCTAATGTT	646		
Qy	1263	TGGCCTCGCTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGTGGTGGA	1322		
Db	647	TGGCCTCGCTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGACCCGTGGTGGA	706		
Qy	1323	CTGACAGTTTCGGAAACCCGGCCGAACCTCGGAGACGTCCTCAGGAGCTTCGGGGGCC	1382		
Db	707	CTGACAGTTTCGAAACCCGGCCGAACCTCGGAGACGTCCTCAGGAGCTTTGGGGGCC	766		
Qy	1383	GTTTTTGTGGCCCGACCTGAGTCTCTAAATCCCGATCGTTTGGAGCTCTTTGGTGCAACC	1442		
Db	767	GTTTTTGTGGCCCGACCTGAGGAGGGAGTCGATGTGAAATCCGACCC-----	814		
Qy	1443	CCCTTAGGAGGAGATATGTGGTCTCGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCCT	1502		
Db	815	-----CGTCAGGATATGTGGTCTCGGTAGGAGACGAGAACCTTAAACAGTTTCCCGCCT	867		
Qy	1503	CCGTCTCAATTTTTCCTTCGGTTTGGACCGAAGCCGCGCGCTCTGTGT--CTGC	1559		
Db	868	CCGTCTCAATTTTTCCTTCGGTTTGGAACCGAAGCCGCGCTCTGTGTCTGTCACGCC	927		
Qy	1560	TGCAGCATCGTTCGTGTTGTCCTGTCGACTGTGTTTCTGTATTTGTCTGAAAAATATG	1619		

Db 1929 CGACCTGACGCAATATGGGATCGCCATTTGAACAGATGGATTGACGCGAGTTCTC-- 1986
Qy 2700 CTTACCATCAGACGCTCCAGCCAGAGGACATCGCCACTACTACTCTGACGCAATATAG 2759
Db 1987 -----CGGCCGCTTTGGGTGGAGAGGCTAT--TCGGCTATGACTGGGCACACAGAC 2035
Qy 2760 CTTATCGTGTGCTCGGCCAAGGACCAAGGTGGAAATCAACAGAGGTGCTCAGGATC 2819
Db 2036 AATCGCTGCTC-----TGATCGCGCGCTGT 2062
Qy 2820 GGCTGGATCCGCGCTCTGCTGGCTCAGGATCGAGTCCAACTGTGTGGAGACGCTGGAGG 2879
Db 2063 CCGGCTGTACGACGAGGGGCGCCGCTTCTTTTGTCAAGACGACCTGTCGGGTGCCCT 2122
Qy 2880 TGTGTGCAACTGCGCCGCTCCCTGCGCTCTGCTGCTCGCATCTGCTGCTGCTGCTGCT 2939
Db 2123 GAATGAACCTGCAGGACGAGGACGCGGCTATCGTGGCTGCGCCACGAGCGGCTTCTTG 2182
Qy 2940 CACATATTTGGATGAGTTGGGTGAGACGACGACCTGGAAAGGTCTTGAGTGAGTTGAGA 2999
Db 2183 CGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTTGGGCGAAGT 2242
Qy 3000 AATTCATCCAGTAGCAGTATGATTAAGTATGCGCC-GTCTCTAAAGATAGATTACAA 3058
Db 2243 GCCGGGCGAGGATCTCTGTCTATCTCCTTGCTCTGCTCCGAGAAAGTATCCATCATGGC 2302
Qy 3059 TATCGCAGACAAAGCCCAAGAACACATTTGTTCTCTGCAATGGACAGCCTGAGACCCGAAG 3118
Db 2303 TGATGCAATCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTTGACACCAAGC 2362
Qy 3119 ACACGGGGTCTATTTTGTGGAAGCTTTACTTGGCTTCCCTGCTTCCCTGCTTATTTGGG 3178
Db 2363 GAAAC--ATCGCATCGAGCGACGACTACTCGGATGGAAGCGCTCTTGTGATCAGGA 2419
Qy 3179 GCCAAGGACCCCGCTCAGCTCTCCAGTGTAAAGCCACACAGCCGACGCGCGGAC 3238
Db 2420 TGATCTGGAAGAGCATCAAGGGCTCGCCGACCGCAATGTTTCGCGGCTCAAGGC 2479
Qy 3239 CACCAACACCGCGCGCCACCATCGGCTCGCAGCCCTGCTGCTGCGCCAGAGGGGCTC 3298
Db 2480 GCGCATGCCGACGGGAGGATCTGCTGCTGACCCATGGC----- 2519
Qy 3299 GCGCAGCGCGGGGGCGCAGTGCAACAGGGGGCTGGACTTCGCGCTGGATCCCAAC 3358
Db 2520 -----GATGCTGCT 2529
Qy 3359 TCTGCTACTGCTGATGGAATCTCTTCATCTATGCTGCTATCTCACTGCTTGTTC 3418
Db 2530 TGCCGAATATCATGTTGGAATGCGCGCTTTTCTGGAATTCATCGACTGTG----- 2580
Qy 3419 TGAGAGTGAAGTTTCAGCAGGAGCGCAGAGCCCCCGGTACAGCAGGCGCCAGAACAGC 3478
Db 2581 -----GCGGCTGGGTGTGGGACC 2601
Qy 3479 TCTATAACAGCTCAATCTAGGACGGAAGAGGAGTACGATGTTTGGACAGAGACGTG 3538
Db 2602 GCTATCAG----- 2610
Qy 3539 GCCGGGACCTGAGATGGGGGGAAGCCGAGAGGAAGAACCTCAGGAAGCCCTGTACA 3598
Db 2611 ----- 2610
Qy 3599 ATGAACTGCAGAAAGATAGATGGGAGGCTACAGTGAGATGGGATGAAGCGGAGC 3658
Db 2611 -----ACATAGGCTTGGCTACCCGCTGATATGCTGAAGAGCTTGGCG 2652
Qy 3659 GCCGAGGGGCAAGGGGACGATGCGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACA 3718
Db 2653 GCGAATGGGTGACCGCTTCTCTGCTTACGGTATCGCCGCTCCCGATGTCAGCGCA 2712
Qy 3719 CTTACGACGCGCTTCAATGACGCGCTTGCCTCTGCTTAAGTGAAGGTTGCTGCGGATC 3778
Db 2713 TCGCCTTCTATCGCCTTCTTGAC----- 2736

Qy 3779 CGGATTAAGTCCAATTTGTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTGACTCA 3838
Db 2737 -----AGTTCTTCTGAGGGGATCGTCCGACCTGCGAGCAAGCTT----- 2775
Qy 3839 ACAATATCACCAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATT 3898
Db 2776 -----ATCGATAAATAAAGATTTTATT 2799
Qy 3899 TAGTCTCCAGAAAAGGGGGAATGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTA 3958
Db 2800 TAGTCTCCAGAAAAGGGGGAATGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTA 2859
Qy 3959 AGTAAACCCATTTTGAAGCATGGAATAATACATACTGAGAAATGAGAGAGTTGAGATCA 4018
Db 2860 AGTAAACCCATTTTGAAGCATGGAATAATACATACTGAGAAATGAGAGAGTTGAGATCA 2919
Qy 4019 AGTTAGGAACAGACAGACAGCAGAAATATGGGCCAACAGGATATCTGTGTAGACAGTT 4078
Db 2920 AGTTAGGAACAGACAGACAGC----- 2941
Qy 4079 CTTGCCCCGCTCAGGGGCAAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATC 4138
Db 2942 -----AGAATATGGGCCAAACAGGATATC 2965
Qy 4139 TGTGTAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCGGTC 4198
Db 2966 TGTGTAAGCAGTTCTGCCCCCGGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCGGTC 3025
Qy 4199 CCGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAA 4257
Db 3026 CCGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAA 3085
Qy 4258 AATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTTCGCGCG 4317
Db 3086 AATGACCTGTGCTTATTTGAACCAATCAGTTCTGCTTCTGCTTCTGTTTCGCGCG 3145
Qy 4318 CTTCTGCTCCCGAGCTCAATAAAGAGCCACAAACCCCTCCTCGGCGGCCAGTCTCT 4377
Db 3146 CTTCTGCTCCCGAGCTCAATAAAGAGCCACAAACCCCTCCTCGGCGGCCAGTCTCT 3205
Qy 4378 CGATAGACTGCTGCTGCGCGGCTACCGTGTCTCAATAAACCCTCTTGAGTTCGATCCG 4437
Db 3206 CGATAGACTGCTGCTGCGCGGCTACCGTGTATCAATAAACCCTCTTGAGTTCGATCCG 3265
Qy 4438 ACTGCTGCTGCTGCTTCTTGGAGGGTCT-CTCTGAGTGAATTGACTACCCGTCAGCG 4496
Db 3266 ACTGCTGCTGCTGCTTCTTGGAGGGTCTCTCTGAGTGAATTGACTACCCGTCAGCG 3325
Qy 4497 GGGTCTTTTCAGTTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTGCCGAGGG 4556
Db 3326 GGGTCTT----- 3333
Qy 4557 ACTCCGTCAGCCCGGTTTTTGTATATAAATGCAAGAACAGTGTTCCTTCAAGCCA 4616
Db 3334 ----- 3333
Qy 4617 GACTACATCTGACTCTCGGCTTTTATAAAGATGTTGAAGGGCTCTGTGAGCTATCTGC 4676
Db 3334 ----- 3333
Qy 4677 CACAGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTCG 4736
Db 3334 ----- 3333
Qy 4737 TCTATTTTGTGGCTTCTCGGTAATTTAAATTTCTAGTTTGCATCTCCCTTCTGAGAGCA 4796
Db 3334 ----- 3333
Qy 4797 CGGCGATTGACAGTAGTTAATCTCTGAGGGCAGGCTTCTGTGAAAGGTTGCTGCGGC 4856
Db 3334 ----- 3333

Db 5462 AAAAGGGAATAAGGCGACACGGAATGTTGAATCTACTACTCTTCTCTTTTCAATAT 5521
 Qy 7076 TATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGATTTAG 7135
 Db 5522 TATTGAAGCATTTATCAGGGTTATTGTCTCATGACGGGATACATATTTGAATGATTTAG 5581
 Qy 7136 AAAATAAACAATAGGGGTTCCGGCGACATTTTCCCGAAAAAGTGCACCTGACGTCTAA 7195
 Db 5582 AAAATAAACAATAGGGGTTCCGGCGACATTTTCCCGAAAAAGTGCACCTGACGTCTAA 5641
 Qy 7196 GAAACCATTTATTCATGACATTAACCTTATAAATAAGCGGTATCACGAGCGCTTTTCGT 7255
 Db 5642 GAAACCATTTATTCATGACATTAACCTTATAAATAAGCGGTATCACGAGCGCTTTTCGT 5701
 Qy 7256 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGCGACTCCCGGAGACGGTC 7315
 Db 5702 CTCGGCGGTTTCGGTGATGACGGTGAAACCTCTGACATGCGACTCCCGGAGACGGTC 5761
 Qy 7316 ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGGTCACGGGT 7375
 Db 5762 ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCGTCAGGGCGGTCACGGGT 5821
 Qy 7376 GTTGGCGGGTTCGGGGTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGTG 7435
 Db 5822 GTTGGCGGGTTCGGGGTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGTG 5881
 Qy 7436 CACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGGCG 7495
 Db 5882 CACCATATGCGGTGTAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGGCG 5941
 Qy 7496 CATTGCGCATTCAGGCTCGCAACTGTTTGGGAAGGCGGATCGTGCGGGCTCTTCGCTA 7555
 Db 5942 CATTGCGCATTCAGGCTCGCAACTGTTTGGGAAGGCGGATCGTGCGGGCTCTTCGCTA 6001
 Qy 7556 TTACGCCAGCTGGCGAAAGGGGATGCTGTCGAAGCGCATTAAGTTGGTTACGCCAGG 7615
 Db 6002 TTACGCCAGCTGGCGAAAGGGGATGCTGTCGAAGCGCATTAAGTTGGTTACGCCAGG 6061
 Qy 7616 TTTTCCAGTCACGACGTTGTAACACGACGCGCAGTGCC 7654
 Db 6062 TTTTCCAGTCACGACGTTGTAACACGACGCGCAGTGCC 6100

RESULT 6
 ID ADQ80674
 XX ADQ80674 standard; DNA; 5903 BP.
 AC ADQ80674;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Bcl-XL expression directing retroviral vector, pFB-bcl-XL.
 XX
 KW Survival; neuron; tyrosine hydroxylase; tyrosine 3-monooxygenase; TH;
 KW anti-apoptotic; Bcl-XL; neurological disorder; neuroprotective;
 KW nootropic; antiparkinsonian; transplantation; drug screening;
 KW gene profiling; CNS disorder; neurodegenerative disease; pFB-bcl-XL; ds.
 XX Unidentified.
 OS
 XX
 XX WO2004062554-A2.
 FN
 XX
 XX 29-JUL-2004.
 PD
 XX
 PF 07-JAN-2004; 2004WO-DK000008.
 XX
 XX 08-JAN-2003; 2003US-0438719P.
 PR 11-APR-2003; 2003DK-00000581.
 PR 22-APR-2003; 2003US-0464546P.
 XX
 XX (NSGE-) NSGENE AS.
 PA
 XX Martinez-Serrano A, Liste I, Villa A;

XX WPI; 2004-544027/52.
 DR
 XX Enhancing the survival of neurons or cells expressing tyrosine
 PT hydroxylase (TH) for treating neurodegenerative disorders, comprises
 PT contacting neurons or TH expressing cells with Bcl-XL or its functional
 PT equivalent.
 XX
 XX Disclosure; SEQ ID NO 5; 108pp; English.
 PS
 XX The invention relates to a novel method for enhancing the survival of
 CC neurons and/or of cells expressing tyrosine hydroxylase (EC 1.14.16.2 -
 CC tyrosine 3-monooxygenase) (TH +). The method comprises contacting a
 CC population of cells with Bcl-XL or its functional equivalent, where the
 CC population of cells is selected from: neurons or cells capable of
 CC differentiating into neurons; or TH expressing cells or cells capable of
 CC differentiating into TH expressing cells. The invention further
 CC comprises: a composition of cells obtainable by the method above; a
 CC composition of isolated mammalian cells overexpressing the anti-apoptotic
 CC Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic
 CC neuron; an implantable cell culture device comprising: a semi-permeable
 CC membrane permitting the diffusion of a biologically active protein
 CC through it; and a composition of cells selected from above; a lentiviral
 CC vector particle being produced based on a lentiviral transfer vector;
 CC enhancing the survival of TH + cells in vivo; a retroviral particle being
 CC produced based on a retroviral transfer vector; enhancing the survival of
 CC in vivo differentiated dopaminergic neurons; a packaging cell line
 CC capable of producing an infective vector particle; a packaging cell line
 CC capable of producing an infective vector particle; treatment of a
 CC neurological disorder; a fusion protein comprising the Bcl-XL sequence
 CC comprising 233 amino acids ADQ80670 or its functional equivalent and a
 CC membrane translocation signal; an expression vector comprising a
 CC polynucleotide sequence coding for the fusion protein and a promoter
 CC sequence capable of directing the expression of the fusion protein in a
 CC host cell; a host cell comprising the expression vector; and producing
 CC the fusion protein. The compositions of the invention have
 CC neuroprotective, nootropic, and antiparkinsonian activities. The cells
 CC are useful for transplantation, drug screening, gene profiling, or for
 CC the preparation of a medicament useful for the treatment of a CNS
 CC disorder. The CNS disorder is a neurodegenerative disease involving
 CC lesioned and traumatic neurons, including traumatic lesions of peripheral
 CC nerves, the medulla, the spinal chord, cerebral ischaemic neuronal
 CC damage, neuropathy, peripheral neuropathy, Alzheimer's disease,
 CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis,
 CC or memory impairment connected to dementia. The method is useful for
 CC enhancing the survival of neurons and/or of cells expressing tyrosine
 CC hydroxylase for the treatment of neurodegenerative disorders. This
 CC sequence represents an expression construct for directing the expression
 CC of Bcl-XL in transduced cells, used in the method for increasing the
 CC survival rate of neurons of the invention.
 XX
 XX Sequence 5903 BP; 1410 A; 1621 C; 1507 G; 1365 T; 0 U; 0 Other;

Query Match 42.1%; Score 3223.6; DB 13; Length 5903;
 Best Local Similarity 73.9%; Pred. No. 0;
 Matches 4964; Conservative 0; Mismatches 759; Indels 995; Gaps 29;
 Qy 550 TTTGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAGGC 609
 Db 167 TGTGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAGGC 226
 Qy 610 ATGG-AAAATACATACTGAGAATAGAGAAGTTCAGATCAAGGTTAGGAACAGA-GAGAC 667
 Db 227 ATGGAATAATACATACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGACAGATGGAC 286
 Qy 668 AGCAGAATATGGGCCAAACAGGATATCTGTGTAAGCAGATTCTCTGCCCC-GCTCAGGGCC 726
 Db 287 AGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGATTCTCTGCCCCGGCTCAGGGCC 346
 Qy 727 AAGAACAGTTTGGAAACAGGAGAAATATGGCCAAACAGGATATCTGTGTAAGCAGATTCTCTG 786
 Db 347 AAGAACAGATGGAACAGCAGCTGAATATGGGCCAAACAGGATATCTGTGTAAGCAGATTCTCTG 406

QY 787 CCCGGCTCAGGCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCTCAGCAGTTTCT 846
 Db 407 CCCGGCTCAGGCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCTCAGCAGTTTCT 466
 QY 847 AGAARACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAATGACCTGTGCTTATT 906
 Db 467 AGAARACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAATGACCTGTGCTTATT 526
 QY 907 TGAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTCTGCGCGCTTCTGCTCCCGAGCTCA 966
 Db 527 TGAACCTAACCAATCAGTTCGCTTCTCGCTTCTGTTCTGCGCGCTTCTGCTCCCGAGCTCA 586
 QY 967 ATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCTCCGATAGACTGCGTCCCGG 1026
 Db 587 ATAAAGAGCCCAACACCCCTCACTCGGGCGCCAGTCTCCGATGACTGAGTCCCGG 646
 QY 1027 GGTACCCGTTATCCCAATTAAGCCTCTTGTGTTTGCATCCGATCCGATCGCTGAT 1086
 Db 647 GGTACCCGTTATCCCAATTAAGCCTCTTGCAG--TTGCATCCGACTTGTGCTCGCTGT 705
 QY 1087 CTTGGGAGGGTCTCCTCAGATTGATTGACTGCC--ACCTCGGGGCTCTTCAATTTGGAG 1145
 Db 706 CTTGGGAGGGTCTCCTCAGATTGACTACCCGTCAGGGGGTCTTTCAATTTGGG 765
 QY 1146 GTTCCACCGAGATTTGGAGACCCCTGCCAGGAGCACCGACCCCGCGGGAGGTAA 1205
 Db 766 GCTCGTCCGGATCGGGAGACCCCTGCCAGGAGCACCGACCCCGAGGGAGTAA 824
 QY 1206 GGTGGCCAGAACTTATCTGTCTGTCTCGATTGTCTAGTGTCTATGACTGATTTTATGC 1265
 Db 825 GGTGGCCAGAACTTATCTGTCTGTCTCGATTGTCTAGTGTCTATGACTGATTTTATGC 884
 QY 1266 GCTCGGTGGTACTAGTACTAGTCTGATCTGATCTGTTCTGCGGACCCGTGTGGAATG 1325
 Db 885 GCTCGGTGGTACTAGTACTAGTCTGATCTGATCTGTTCTGCGGACCCGTGTGGAATG 944
 QY 1326 ACAGTTTCGGAAACACCGCGCCGCAACCTCTGGAGACGTCCTCAGGAGCTTCGCGGGCGGT 1385
 Db 945 ACAGTTTCGGAAACACCGCGCCGCAACCTCTGGAGACGTCCTCAGGAGCTTCGCGGGCGGT 1004
 QY 1386 TTTGTGGCCGACCTGAGTCTTAAATCCGATCGTTTATAGGACTCTTTGGTGACACCCCG 1445
 Db 1005 TTTGTGGCCGACCTGAGTCTTAAATCCGATCGTTTATAGGACTCTTTGGTGACACCCCG 1064
 QY 1446 TTAGAGGGGATATGTTGTTCTGTTAGGACGAGAACTTAAACAGTTCGCGCTCG 1505
 Db 1065 TTAGAGGGGATATGTTGTTCTGTTAGGACGAGAACTTAAACAGTTCGCGCTCG 1124
 QY 1506 TCTGAATTTTGTCTTTCGGTTTGGACCGACCGCGCGCTCTTGTCTGCTGCAGC 1565
 Db 1125 TCTGAATTTTGTCTTTCGGTTTGGACCGACCGCGCGCTCTTGTCTGCTGCAGC 1184
 QY 1566 ATGTTCTGTGTCTGTCTGACTGTGTTTCTGTAATTTGCTGAAATATGGGCCG 1625
 Db 1185 ATGTTCTGTGTCTGTCTGACTGTGTTTCTGTAATTTGCTGAAATATGGGCCG 1244
 QY 1626 GGTAGACTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACTGNAAGATGTCAGCGG 1685
 Db 1245 GGTAGACTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACTGNAAGATGTCAGCGG 1304
 QY 1686 ATCGCTCACAACAGTTCGGTATGATGTCAGAGAGACGTTGGTTACCTTCTGCTGCA 1745
 Db 1305 ATCGCTCACAACAGTTCGGTATGATGTCAGAGAGACGTTGGTTACCTTCTGCTGCA 1364
 QY 1746 GAATGGCCAACTTTAACTCGGATGGCCCGGACCGGACCTTTTAAACGAGACCTTATC 1805
 Db 1365 GAATGGCCAACTTTAACTCGGATGGCCCGGACCGGACCTTTTAAACGAGACCTTATC 1424
 QY 1806 ACCAGGTTAAGATCAAGTCTTTTCACTGGCCCGCATGGACACCCAGACAGGTCCCG 1865
 Db 1425 ACCAGGTTAAGATCAAGTCTTTTCACTGGCCCGCATGGACACCCAGACAGGTCCCG 1484
 QY 1866 TACATCGTACCTGGGAAGCCTTGGCTTTTGTAGCCCGCTCCCTCGGTCAAGCCCTTGTGA 1925

Db 1485 TACATCGTACCTGGGAAGCCTTGGCTTTTGTAGCCCGCTCCCTGGGTCAAGCCCTTGTGA 1544
 QY 1926 CACCCCTAAGCCTCCGCTCTCTTCTCCATCCGCGCGGTCTCTCCCTTGAACCTCTCT 1985
 Db 1545 CACCCCTAAGCCTCCGCTCTCTTCTCCATCCGCGCGGTCTCTCCCTTGAACCTCTCT 1604
 QY 1986 CQTTCGACCCCGCTCGATCCTCTTATCCAGCCCTCAGTCTCTCTTAGCGCGCCCG 2045
 Db 1605 CQTTCGACCCCGCTCGATCCTCTTATCCAGCCCTCAGTCTCTCTTAGCGCGCCCG 1664
 QY 2046 ATATGGCCATATGAGATCTTATATGGGCAACCCCGCGCTTGTAACTTCCCTGACCT 2105
 Db 1665 ATATGGCCATATGAGATCTTATATGGGCAACCCCGCGCTTGTAACTTCCCTGACCT 1724
 QY 2106 GACATGACAGAGTTTACTTAAAGCCCTCTCTCCAGCTCAGTCTCTTACTTCTT 2164
 Db 1725 GACATGACAGAGTTTACTTAAAGCCCTCTCTCCAGCTCAGTCTCTTACTTCTT 1784
 QY 2165 GTCCAGCAGAACTCTGGAGACTCTGGCGGAGCTTACCAAGAACTTACCAAGAACT 2224
 Db 1785 GTCCAGCAGAACTCTGGAGACTCTGGCGGAGCTTACCAAGAACTTACCAAGAACT 1843
 QY 2225 GTGTACCTTACCTTACCGAGTCCGCGACACAGTGTGGTCCGCGACACAGACTAAG 2284
 Db 1844 GTGTACCTTACCTTACCGAGTCCGCGACACAGTGTGGTCCGCGACACAGACTAAG 1903
 QY 2285 AACCTAGAACCTCGCTGGAAGGACTTACACAGTCTCTGCTGACACCCCGACCCGCTC 2344
 Db 1904 AACCTAGAACCTCGCTGGAAGGACTTACACAGTCTCTGCTGACACCCCGACCCGCTC 1963
 QY 2345 AAGTAGACGGCATCGAGCTTGGATACAGCGCGCCACGCTGAAGGCTGCGACCCCGG 2404
 Db 1964 AAGTAGACGGCATCGAGCTTGGATACAGCGCGCCACGCTGAAGGCTGCGACCCCGG 2023
 QY 2405 GGTGAGCACTCTCTAGACTGCGATGGGATGGAGCTGTATCATCTCTTCTTGGTAGCAA 2464
 Db 2024 GGTGAGCACTCTCTAGACTGCGATGGGATGGGATGGGATGGGATGGGATGGGATGG 2082
 QY 2465 CAGCTACAGGTGTCTCTCGACATCCAGTGTGAGCCAGAGCCAGAGCCAGCTGAGGCCA 2524
 Db 2083 CCGCGCCGAGTGTCTGAGGAGCCATCTCTATTTATATAA--TGTCTCAGAGCAA 2134
 QY 2525 CGTGGGTGACAGAGTGTGACCTCTGTAGGCGCAGTCTGAGTGTGGTACTTCTGTAG 2584
 Db 2135 CCGGAGAGTGTGTGTGTCTTCTCTTAAGCTCTCCAGAGAGATACAGTGTGAG 2194
 QY 2585 CTTGTGATCCAGCAGAGGAGGATGAGGCTCCTGATCTGATCTGATCTGATCTGATCTG 2644
 Db 2195 TCAGTTTAGCGATGTGGAAGAGAACAGGACTGAGCCCGCAGAGAACTGAAACCAAGAA 2254
 QY 2645 GGCACTGTGTGTGCAAGAGAGATTCAGGAGTGTGAGGTAGCGGTACCGACTTCACTTCA 2704
 Db 2255 GGAGACCCCGAGTGTGATCAATGGCAACCCATCTTGGCAGC--TGGCGATAGCCCGCG 2312
 QY 2705 CCATCAGCAGCTCCAGCGCAGAGGACATCGCCACCTACTCTGCGCAGCAATATAGCTCT 2764
 Db 2313 GTGAATGAGGACCTTGGCCACAG-----CAGCAGTTTGGATGGC 2351
 QY 2765 ATCGGTCTGTCGGCCCAAGGACCAAGGTGGAATCAAAAGAGGTGGCTCAGGATCGGCTG 2824
 Db 2352 CCGGAGGTACTTCCCATGTGCAGAGTGAAGCAAGCGCTGAGAGAGGTGGCGAT----- 2405
 QY 2825 GATCCGGTCTGGTGGCTCAGGATCGGAGGTCTCACTGTGGTGGAGAGCGGTGAGGTGTTG 2884
 Db 2406 -----GAGTTTGAATCTCGGTACCGGAGAGCATTCAGGTG 2439
 QY 2885 TGCAACCTGGCGGTCCCTGCGCTGTCTGTCTCCGCTCTGTGCTTCCGATTTTCAACCAT 2944
 Db 2440 ATCTAACATCCAG-----CAGTTTGAATCTCGGTACCGGAGAGCATTCAGGTG 2459
 QY 2945 ATTGGATGAGTTGGGTGAGACAGGCACTTGGAAAAAGGTCTTGTGAGTGGATTTGAGAAATTC 3004

Db 2460 ATAAACCCAGGGACAGTATATCAGAGCTTTGAAACAGGTAGTG----- 2501
Qy 3005 ATCCAGATAGCAGTACGATTAACATATATGCGCGCTCTTAAGAGTAGATTTTACAATATCGC 3064
Db 2502 ----- 2501
Qy 3065 GAGACACGCCAGAAACACATTTGTCCTGCAAAATGACAGCCTGAGACCCCGAAGACACCG 3124
Db 2502 -----AATGAACCTCTTTTCGGGATGGGTAAACTG 2530
Qy 3125 GGGTCTAATTTTGTCAAGCCTTTTACTTCGGCTTCCCTGGTTTCTTTATTGGGGCCAAG 3184
Db 2531 GGGTCGCAATGT----- 2542
Qy 3185 GGACCCCGGTCAACCGTCTCCAGTGCTAAGCCCAACACGACGCGCGGACCAACCA 3244
Db 2543 ----- 2542
Qy 3245 CACGGCGCCACCATCGGTGCGAGCCCTGTCTCCCTGGCCCGCAGAGCGGGTGGGCCAG 3304
Db 2543 -----GGCCTCTCTCTTTGGGGGGCACTGTGCGTGAAGAGCGTAGACAAGGAGATGC 2598
Qy 3305 CGGCGGGGGCGCAGTGCACAGAGGGGCTGGACTTCGCCCTGATCCCAAACTCTGCT 3364
Db 2599 AGGTATTGGTGAATGCGGATTCGAAGTTGGATGGCCACCTACTGAATGACCACCTAGAGC 2658
Qy 3365 ACCTGCTGGATGGAACTCTTTCATCTATGTGTCAATTCCTCACTGCCCTGTTCCTGAGAG 3424
Db 2659 CTGTGATCCAGGAGAACGGC-----GGCTGGGACACTTTTG 2694
Qy 3425 TGAAGTTACAGAGAGCCGACAGAGCCCGCGGTACACAGAGGGCGAGAACACAGCTCTATA 3484
Db 2695 TGGATCTCTACGGGAACAAATACAGCACCCGAGAGCCGGAAGGCGCAGGAGGTTTCAACC 2754
Qy 3485 ACGAGCTCAATCTAGGACGAGAGAGGAGTACGATGTTTTGGACAGAGAGCTGGCCGGG 3544
Db 2755 GCTGGTCTCAGCGGGCATGACTGTGGCTGTGTAGTTCTG----- 2795
Qy 3545 ACCCTGAGATGGGGGAAAGCCGAGAGGAAGAACCCCTCAGGAGGCCGTGTACAATGAAC 3604
Db 2796 ----- 2795
Qy 3605 TGCAGAAAGATAAGATGGCGGAGGCTTACAGTAGATTTGGATGAAGGCGAGCGCGGA 3664
Db 2796 ----- 2795
Qy 3665 GGGCAGGGGSCAGATGGCCCTTTACCAGGGTCTCAGTACAGCCACCAGGACACCTACG 3724
Db 2796 ----- 2795
Qy 3725 ACGCCCTTCACATGCAAGCCCTGCGCCCTCGCTAACTCGAGCGCGCGGATCGGAT 3784
Db 2796 -----CTGGGCTCACTCTTCAGTCGGAAGTGAACAGATT 2829
Qy 3785 AGTCAAATTTGTTAAAGACAGATATCAGTGGTCCAGGCTCTAGTCTTTTGACTCAACAATA 3844
Db 2830 CTGCAGATATCCATCACACTGCGCGCGGAATTCGGGGCGCGCATCCGAGTCTTCTGA 2889
Qy 3845 TCACCAGCTGAAGCCTATAGAGTAGGACCATAGATAAAATAAAGATTTTATTAGTCT 3904
Db 2890 GCG-----GGACTCTGGGGTTCGATRAAAATAAAGATTTTATTAGTCT 2933
Qy 3905 CCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTCGCAAGCTAGCTTAAGTAAC 3964
Db 2934 CCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTCGCAAGCTAGCTTAAGTAAC 2993
Qy 3965 GCCATTTTGAAGGCATGG-AAAAATACATAACTGAGATAGAGAAGTTTCAGATCAAGGTT 4023
Db 2994 GCCATTTTGAAGGCATGGAAAATAATACATACTGAGATAGAGAAGTTTCAGATCAAGGTC 3053
Qy 4024 AGGAACAGA-GAGACAGCAGATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTCTCTG 4082
Db 3054 AGGAACAGATGGAACAGCTGAATATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTCTCTG 3113

Qy 4083 CCCC-GCTCAGGGCCAGACACAGTTGGAAACAGGAGATATGGGCCAAACAGGATATCTGT 4141
Db 3114 CCCCAGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGT 3173
Qy 4142 GGTAAAGAGTTCTGCGCCCGGCTCAGGGCCAGAAACAGATGTTGCCAGAGTCGGTCCCG 4201
Db 3174 GGTAAAGCAGTTCTGCGCCCGGCTCAGGGCCAGAAACAGATGTTGCCAGATGCGGTCCAG 3233
Qy 4202 CCCTCAGCAGTTCTTAGAAGAACCATCAGATGTTTCCAGGGTCCCAAGGACCTCAAAATG 4261
Db 3234 CCCTCAGCAGTTCTTAGAAGAACCATCAGATGTTTCCAGGGTCCCAAGGACCTCAAAATG 3293
Qy 4262 ACCCTGTGCTTATTTTGAACATAACCAATCAGTTGCTTCTGCTTCTGCTGCGGCTTC 4321
Db 3294 ACCCTGTGCTTATTTTGAACATAACCAATCAGTTGCTTCTGCTTCTGCTGCGGCTTC 3353
Qy 4322 TGCTCCCGAGCTCAATAAAAGAGCCCAACCCCTCACTCGGCGGCGCAGTCTCCGAT 4381
Db 3354 TGCTCCCGAGCTCAATAAAAGAGCCCAACCCCTCACTCGGCGGCGCAGTCTCCGAT 3413
Qy 4382 AGACTGGTCCCGGGTACCCTGTTCTCAATAAACCTCTTGAGTTGATCCGACTC 4441
Db 3414 TGACTGATGCCCGGGTACCCTGTTATCCAAATAAACCTCTTGAGTTGATCCGACTC 3473
Qy 4442 GTGGTCTGCTGTTCTTGGGAGGGTCT-CTCTGAGTGATTGACTTACCCGTCAGCGGGT 4500
Db 3474 GTGGTCTGCTGTTCTTGGGAGGGTCTCTCTGAGTGATTGACTTACCCGTCAGCGGG 3533
Qy 4501 CTTTCAGTTTCTCCCACTACACAGGTTCTACTAACTCTCTGATGTCGCGGAGGACTC 4560
Db 3534 TCTTTTCAATTTGGGGCTCGTCCGGGATCGGAGACCCCTGCGGAGACCAACCGACCCAC 3593
Qy 4561 CGTCAGCCCGGTTTTTTTATATAATAAATCAAGACAGTGTCCCTTCAAGCCAGACT 4620
Db 3594 CACCGGAGGTAAGCTG----- 3610
Qy 4621 ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACTATCTGCCACA 4680
Db 3611 ----GCTGCTCGCGGCTTTCGGTGATGACGGTGAACCTCTG-ACACATGACGCTCC 3665
Qy 4681 CGATTTTTTAAGATTTTATGCTCTCGGATGAGGGATTTAGTCAATCTATCTCTGCTA 4740
Db 3666 GGAGACGCTCACAGCTGTCTGTAAGCGGATGCGGGAGCAGACAAGCCGTCAGGGCGC 3725
Qy 4741 TTTTGTGCTGCTCTCGTATTTTAAATTTCTAGTTTGACTCCCTTCTGAGACACGCG 4800
Db 3726 GTCAGCGGGTGTGGCG----- 3742
Qy 4801 GATTGCAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGGGCTCAG 4860
Db 3743 ----- 3742
Qy 4861 TGTGAGATTTTGCATAAAAGAGGGTCTCGCCCTGTGTACAGACAGATCGGAATCTAGA 4920
Db 3743 ----- 3742
Qy 4921 GTGCATACTCAGAGTCCCCGGGTTTCGGGCTCTGATCTCAGGGCATCTTTGCTAGAG 4980
Db 3743 -----GGTGTGCGGGCGAGCCATGA----- 3763
Qy 4981 ATCTCTACGCGGACGATCGTGGCGGGTACCAGAGCTCGAAATTCGTAATCATGGTCA 5040
Db 3764 ----- 3763
Qy 5041 AGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCCAACAACATACGAGCCGAA 5100
Db 3764 -----CCAGTCACGCTAGCGATAGCGGA 3786
Qy 5101 GCATAAAGTGAAGCCTGGGGTCCCTAATGAGTGAAGTGAAGTCACTCACTAATTTGGTTGC 5160
Db 3787 GTGTATATCGCTTAACCTATATGCGGCATCAGAGCAGATTTGTACTGAGAGT----- 3835

QY 5161 GCTCACTGCCGCTTTCCAGTCGGGAAACCTGTCGTCGACGCTGCATTAATGAATCGGCC 5220
 Db 3836 -----GCACCATATGCGGTGTGAATAACCGC 3861
 QY 5221 AACCGCGGGAGAGCGGTTTGGGTATTGGGGCTCTTCGGTTCCTCGCTCACTGACT 5280
 Db 3862 ACAGATGCGTAAGGAGAAATACCGCATCAGGCGCTCTTCGGTTCCTCGCTCACTGACT 3921
 QY 5281 CGTCGCGCTCGGTTCGGCTGCGCGAGCGGTATCAGTCACTCAAAAGGGGGTAATAC 5340
 Db 3922 CGTCGCGCTCGGTTCGGCTGCGCGAGCGGTATCAGTCACTCAAAAGGGGGTAATAC 3981
 QY 5341 GGTATTCCACAGAAATCAGGGGTAACCGCAGGAAAGAAATGTGAGCAAAAAGGCGACAAA 5400
 Db 3982 GGTATTCCACAGAAATCAGGGGTAACCGCAGGAAAGAAATGTGAGCAAAAAGGCGACAAA 4041
 QY 5401 AGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTCCTCAATAGGCTCCGCGCCCTCG 5460
 Db 4042 AGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTCCTCAATAGGCTCCGCGCCCTCG 4101
 QY 5461 ACAGCATCACAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCGCAGGACTATAA 5520
 Db 4102 ACAGCATCACAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACCGCAGGACTATAA 4161
 QY 5521 GATACAGGCGTTTCCCGCTTCCGCTTCCGCTCTCCTGTTCCGAACCGCTGCCGC 5580
 Db 4162 GATACAGGCGTTTCCCGCTTCCGCTTCCGCTCTCCTGTTCCGAACCGCTGCCGC 4221
 QY 5581 TTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAC 5640
 Db 4222 TTACCGGATACCTGTCGCGCTTCTCCCTTCGGAAGCGTGGCGCTTCTCATAGCTCAC 4281
 QY 5641 GCTGTAGGTATCTCAGTTCGCTGAGTGTGCTCCAGCTGGCGTGTGTCAGCAAC 5700
 Db 4282 GCTGTAGGTATCTCAGTTCGCTGAGTGTGCTCCAGCTGGCGTGTGTCAGCAAC 4341
 QY 5701 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 5760
 Db 4342 CCCCCCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGG 4401
 QY 5761 TAAGACAGACTTATCGCACTGGCAGAGCACTGGTAAACAGGATTAGCAGAGCGAGT 5820
 Db 4402 TAAGACAGACTTATCGCACTGGCAGAGCACTGGTAAACAGGATTAGCAGAGCGAGT 4461
 QY 5821 ATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGCTTAACCTACGGCTACCTAGAGGA 5880
 Db 4462 ATGTAGCGGTGCTACAGAGTCTTCAAGTGTGGCTTAACCTACGGCTACCTAGAGGA 4521
 QY 5881 CAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT 5940
 Db 4522 CAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCT 4581
 QY 5941 CTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGAGA 6000
 Db 4582 CTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGAGA 4641
 QY 6001 TTACGGCAGAAAAAGGATCTCAGAGAGTCTTTGATCTTTTCTACGGGTCTGACG 6060
 Db 4642 TTACGGCAGAAAAAGGATCTCAGAGAGTCTTTGATCTTTTCTACGGGTCTGACG 4701
 QY 6061 CTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 6120
 Db 4702 CTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTATCAAAAAGGATCT 4761
 QY 6121 TCACCTAGATCCTTTTAAATTAAGTGTGTTTTTAAATCAATCTAAAGTATATAGT 6180
 Db 4762 TCACCTAGATCCTTTTAAATTAAGTGTGTTTTTAAATCAATCTAAAGTATATAGT 4821
 QY 6181 AACTTTGGTCTGACAGTACCAGTCTTAATCAGTGAGGCACTATCTCAGCGATCTGTC 6240
 Db 4822 AACTTTGGTCTGACAGTACCAGTCTTAATCAGTGAGGCACTATCTCAGCGATCTGTC 4881
 QY 6241 TATTTCTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAAATACGAGGAGG 6300

Db 4882 TATTTCTGTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGAGG 4941
 QY 6301 GCTTACCATCTGGCCCGCAGTGTCAATGATACCGGAGACCCACGCTCACCGGCTCCAG 6360
 Db 4942 GCTTACCATCTGGCCCGCAGTGTCAATGATACCGGAGACCCACGCTCACCGGCTCCAG 5001
 QY 6361 ATTTATCAGCAATAAACCAGCAGCGGAGGCGGAGCGCAGAGAGTGGTCTCTGCAACTT 6420
 Db 5002 ATTTATCAGCAATAAACCAGCAGCGGAGGCGGAGCGCAGAGAGTGGTCTCTGCAACTT 5061
 QY 6421 TATCCGCTTCCATCCAGTCTTAAATGTTGCGGGAAGCTAGAGTAAGTGTTCGCCAG 6480
 Db 5062 TATCCGCTTCCATCCAGTCTTAAATGTTGCGGGAAGCTAGAGTAAGTGTTCGCCAG 5121
 QY 6481 TTAATAGTTTGGCAACCGTTGTCATTCGTCGAGGCAATCGTGTGTCCGCTCGTCT 6539
 Db 5122 TTAATAGTTTGGCAACCGTTGTCATTCGTCGAGGCAATCGTGTGTCCGCTCGTCT 5181
 QY 6540 TTGGTATGGCTTCATTCAGTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCCA 6599
 Db 5182 TTGGTATGGCTTCATTCAGTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCCA 5241
 QY 6600 TGTGTGCAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTGTGTCAGAACTAAGTTGG 6659
 Db 5242 TGTGTGCAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTGTGTCAGAACTAAGTTGG 5301
 QY 6660 CCGCAGTGTATCAGTCACTGTTATGCGAGCACTGCATAATCTCTTACTGTCAGCCAT 6719
 Db 5302 CCGCAGTGTATCAGTCACTGTTATGCGAGCACTGCATAATCTCTTACTGTCAGCCAT 5361
 QY 6720 CCGTAAAGTGTCTTTCTGACTGCTGAGTACTCAACCAAGTCACTCTGAGAAATAGTGA 6779
 Db 5362 CCGTAAAGTGTCTTTCTGACTGCTGAGTACTCAACCAAGTCACTCTGAGAAATAGTGA 5421
 QY 6780 TCGCGCAGCCGAGTGTCTTTGCGCGGCTCAATACGGGATTAATACCGCGCACATGCA 6839
 Db 5422 TCGCGCAGCCGAGTGTCTTTGCGCGGCTCAACACGGGATAATACCGCGCACATGCA 5481
 QY 6840 GAACTTTAAAGTGTCTCATCATGAAACCGTTCTTCGGGGGAAAACCTCTCAAGGATCT 6899
 Db 5482 GAACTTTAAAGTGTCTCATCATGAAACCGTTCTTCGGGGGAAAACCTCTCAAGGATCT 5541
 QY 6900 TACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCTGTCACCCCAACTGATCTTCAGCAT 6959
 Db 5542 TACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCTGTCACCCCAACTGATCTTCAGCAT 5601
 QY 6960 CTTTACTTTTCCACCGCTTCTGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA 7019
 Db 5602 CTTTACTTTTCCACCGCTTCTGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA 5661
 QY 7020 AGGGAATAAGGGCGACACGGAATGTTGAATCTCATACTCTCTCTTTTCAATATTAT 7079
 Db 5662 AGGGAATAAGGGCGACACGGAATGTTGAATCTCATACTCTCTCTTTTCAATATTAT 5721
 QY 7080 GAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTTATTAGAAA 7139
 Db 5722 GAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTTATTAGAAA 5781
 QY 7140 ATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTCTAAGAAA 7199
 Db 5782 ATAAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTCTAAGAAA 5841
 QY 7200 CCAATTATTCATGATTAACCTATATAAATAGGCGTATACGAGGCGCTTTTCGTCT 7257
 Db 5842 CCAATTATTCATGATTAACCTATATAAATAGGCGTATACGAGGCGCTTTTCGTCT 5899

RESULT 7
 ABR85887
 ID ABK85887 standard; DNA; 5782 BP.
 XX
 AC ABK85887;

	XX DT	21-AUG-2002 (first entry)		Qy	789	CCGGCTCAGGCGCAAGAACAGATGGTGTCCCAGATGCGGTCCCGCCCTCAGCAGTTTCCTAG	84
	XX DE	DNA sequence encoding Gag packaging protein.		Db	243	CCGGCTCAGGCGCAAGAACAGATGGTGTCCCAGATGCGGTCCCGCCCTCAGCAGTTTCCTAG	302
	XX KW	Gag; ds; viral expression vector.		Qy	849	AGAACCATCAGATGTTTTCCAGGGTGCCCAAGACACTGAATAACACCTGTGCCTTAATTG	908
	XX OS	Unidentified.		Db	303	AGAACCATCAGATGTTTTCCAGGGTGCCCAAGACACTGAATAACACCTGTGCCTTAATTG	362
	XX PN	WO200234929-A2.		Qy	909	AACCTAACAATCAGTTTCGGCTTCTCGCTTCTGTGTCGCGCGCTTCTGCTCCCCGAGCTCAAT	968
	XX PD	02-MAY-2002.		Db	363	AACCTAACAATCAGTTTCGGCTTCTCGCTTCTGTGTCGCGCGCTTCTGCTCCCCGAGCTCAAT	422
	XX XX	18-OCT-2001; 2001WO-US032592.		Qy	969	AAAGAGCCCAACACCCCTCATCTCGGCGGCCAGTCTCTCGATAGACTGTGCTGCCCGGG	102
	XX PR	20-OCT-2000; 2000US-0241879P.		Db	423	AAAGAGCCCAACACCCCTCATCTCGGCGGCCAGTCTCTCGATAGACTGTGCTGCCCGGG	482
	XX PA	(WHED) WHITEHEAD INST BIOMEDICAL RES.		Qy	1029	TACCCTGATTTCCCAATAAAGACCTTCTGTGTTTGCAATCCGAATCGTGGACTCGCTGATCC	108
	XX PI	Kohn BY, Daley GQ;		Db	483	TACCCTGATTTCCCAATAAAGACCTTCTGTGAG - TTGCATCCGACTTGTGGTCTCGCTGTTCC	544
	XX DR	WPI; 2002-489949/52.		Qy	1089	TTGGAGGGTCTCTCTCAGATTGATTGACTGCCCC - ACTCTGGGGGGTCTTTTCATTGGAGGT	114
	XX PT	Expression vectors comprising viral vectors (e.g. retroviral vectors), useful for generating or screening nucleic acid libraries, or in antisense based or gene trapping methods for identifying modulators of a mammalian gene.		Db	542	TTGGAGGGTCTCTCTCAGATTGATTGACTACCGTCAAGCGGGGGTCTTTTCATTGGGGGC	601
	XX PS	Claim 8; Fig 1; 114pp; English.		Qy	1148	TCACACGAGATTTGGAGACCCCTGCCACGAGCACACCGACCCCGCGGGAGGTAAGC	120
	XX CC	This invention relates to a novel vector comprising from 5' to 3' a packaging sequence, a heterologous insert sequence or restriction site for insertion of a heterologous sequence and a 3' long terminal repeat (LTR) sequence, where at least 2 codons of the packaging sequence are altered to reduce formation of fusion polypeptides encoded by the packaging sequence or a portion of it, and the heterologous insert sequence. The vector of the invention is useful for generating a library and in screening nucleic acid libraries. In particular, the vector is useful for identifying and isolating insert nucleic acids based upon their ability to complement a mammalian cellular phenotype in antisense based methods for identifying and isolating nucleic acids that inhibit or reduce function of a mammalian gene or in gene trapping methods to identify and isolate mammalian genes that are modulated in response to a specific stimuli. The vector of the invention exhibits improvements and advantages over prior retroviral expression cloning systems. The vector provides high viral titers to facilitate screening of a complete set of independent cDNAs and high levels of gene expression and ease of recovery of the desired cDNA. The present sequence represents a gag packaging vector of the invention		Db	602	TCGTCCGGATCGGAGACCCCTGCCACGAGCACACCGA - CCGACACCGGAGGTAAGC	660
	XX CC			Qy	1208	TGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTTTATGCCG	126
	XX CC			Db	661	TGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTTTATGCCG	720
	XX CC			Qy	1268	CTGGCTCGGTACTAGTTAGCTAACTAGTCTGTATCTGGCGGACCCGCTGGTAGACTGAACTGAC	132
	XX CC			Db	721	CTGGCTCGGTACTAGTTAGCTAACTAGTCTGTATCTGGCGGACCCGCTGGTAGACTGAACTGAC	780
	XX CC			Qy	1328	GAGTTGGGAACACCCGCGCGCAACCTCTGGAGAGCTGCCAGGACTTCCGGGGGCGCTTTT	138
	XX CC			Db	781	GAGTTGGGAACACCCGCGCGCAACCTCTGGAGAGCTGCCAGG-----TCGGGGGCGCGCTTTT	836
	XX CC			Qy	1388	TGTGGCCGACCTGAGTCTCTAAAATCCGATCCGATCGTTTAGGACTTTTGGTGGCAACCCCTT	144
	XX CC			Db	837	TGTGGCCGACCTGAGTCTCTAAAATCCGATCCGATCGTTTAGGACTTTTGGTGGCAACCCCTT	896
	XX CC			Qy	1448	AGAGGAGGGATATGTGGTCTCTGGTAGGAGACGAGAACCTAAAAACAGTTCCTCCGCTCCGTC	150
	XX CC			Db	897	AGAGGAGGGATATGTGGTCTCTGGTAGGAGACGAGAACCTAAAAACAGTTCCTCCGCTCCGTC	956
	XX CC			Qy	1508	TGAATTTTTCGCTTTCGGTTTGGACACGAGCCGCGCGCGCTTGTCTGTCTGCAGCAT	156
	XX CC			Db	957	TGAATTTTTCGCTTTCGGTTTGGACACGAGCCGCGCGCGCTTGTCTGTCTGCAGCAT	101
	XX CC			Qy	1568	CGTTCTGTGTGTCTCTGTCTGACTGTGTGTCTGTATTTTGTCTGAAAAATATGGGCCCGGG	162
	XX CC			Db	1017	CGTTCTGTGTGTCTCTGTCTGACTGTGTGTCTGTATTTTGTCTGAAAAATATGGGCCCGGG	107
	XX Qy			Qy	1628	CTAGACTGTTTACCACCTCCCTTAAGTTTACCTTTAGTCTCACTGGAAGAAGTGTGAGCGGAT	168
	XX Db			Db	1077	CCAGACTGTTTACCACCTCCCTTAAGTTTGTGACTTGTAGTCTCACTGGAAGAAGTGTGAGCGGAT	113
	XX Qy			Qy	1688	CGCTCAACACAGTTCGGTAGATGTCAAGAAGAGACGTTGGGTTTACCTTCTGTCTGCAGA	174
	XX Db			Db	1137	CGCTCAACACAGTTCGGTAGATGTCAAGAAGAGACGTTGGGTTTACCTTCTGTCTGCAGA	119
	XX Qy			Qy	1748	ATGGCCAACTTTAAAGTTCGGATGGCGCGGAGACGCGACTTTTAAACCGGAGACTCATAC	180
	XX Db			Db	1197	ATGGCCAACTTTTAAAGTTCGGATGGCGCGGAGACGCGACTTTTAAACCGGAGACTCATAC	125
	XX Qy			Qy	1808	CCAGGTTTAAAGTCAAGTCTTTTCACTGCGCCGATGACACCCAGACAGAGTCCCTTA	186
	XX Db			Db	1257	CCAGGTTTAAAGTCAAGTCTTTTCACTGCGCCGATGACACCCAGACAGAGTCCCTTA	131
	XX Qy			Qy	1868	CATCGTACCTCGGGAAGCCTTTGGCTTTTGAACCCCTCCCTCGGTCAAGCCCTTTGTACA	192
	Query Match	42.0%; Score 3213; DB 6; Length 5782;		SQ			
	Best Local Similarity	73.3%; Pred. No. 0;					
	Matches 4860; Conservative	0; Mismatches 905; Indels 864; Gaps 24;					
				Qy	552	TGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCGCCATTTTCAAGGCAT	611
				Db	3	TGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCGCCATTTTCAAGGCAT	62
				Qy	612	GG-AAAAATACATACTCAGATAAGAAAGTTTCAAGTCAAGGTTAGGAACAGA-GAGACAG	669
				Db	63	GGAAAAATACATACTCAGATAAGAAAGTTTCAAGTCAAGGTTAGGAACAGA-GAGACAG	122
				Qy	6		

Db	1317	CATCGTACCTGGGAAGCCTTGGCTTTTGACCCCTCCCTGGGTCAAGCCCTTTGTACA	1376
Qy	1928	CCCTAAGCCCTCGCCCTCTCTTCCTCCATCCGCGCCGCTCTCTCCGCTTGAACCTCTCG	1987
Db	1377	CCCTAAGCCCTCGCCCTCTCTTCCTCCATCCGCGCCGCTCTCTCCGCTTGAACCTCTCG	1436
Qy	1988	TTTCGACCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCCGCCCAT	2047
Db	1437	TTTCGACCCGCTCGATCCTCCCTTTATCCAGCCCTCACTCTCTCTAGGGCCGCCCAT	1496
Qy	2048	ATGGCCATATGAGATCTTATATGGGACACCCCGCCCTTGTAAACTTCCCTGACCCCTGA	2107
Db	1497	ATGGCCATATGAGATCTTATATGGGACACCCCGCCCTTGTAAACTTCCCTGACCCCTGA	1556
Qy	2108	CATGACAAGAGTTACTAAGAGCCCTCTCTCAAGCTCACTTACAGGCT-TCTACTTAGT	2166
Db	1557	CATGACAAGAGTTACTAAGAGCCCTCTCTCAAGCTCACTTACAGGCTCTCTACTTAGT	1616
Qy	2167	CCAGCACGAAGTCTGGAGACCTCTGGCGCAGCCTACCAAGAACAACTGGACCGACCGGT	2226
Db	1617	CCAGCACGAAGTCTGGAGACCTCTGGCGCAGCCTACCAAGAACAACTGGACCGACCGGT	1676
Qy	2227	GGTACCTCAACCTTACCGAGTCGGGACACAGTGTGGTCCGCGACACAGACTAAGAA	2286
Db	1677	GGTACCTCAACCTTACCGAGTCGGGACACAGTGTGGTCCGCGACACAGACTAAGAA	1736
Qy	2287	CCTAGAACCTCGTGGTGAAGAGCTTACAGTCTCTGTGACACACCCCGCCCTCAA	2346
Db	1737	CCTAGAACCTCGTGGTGAAGAGCTTACAGTCTCTGTGACACACCCCGCCCTCAA	1796
Qy	2347	AGTAGAGCATCGAGCTTGATACACGCCGCCACGCTGAAGGCTGCCGACCCCGGGGG	2406
Db	1797	AGTAGAGCATCGAGCTTGATACACGCCGCCACGCTGAAGGCTGCCGACCCCGGGGG	1856
Qy	2407	TGGACCATCTCTAGACTGCCATGGAGTGTATCATCTCTTTTGGTGTAGCAACA	2466
Db	1857	TGGACCATCTCTAGACTGCCGATCCAGTGTGGTGTAGGAATTCAGCTTGTATCTC	1916
Qy	2467	GCTACAGGTGTCCATCCGACATCCAGCTGACCCGAGCCCAAGAGCCTGAGCGCCAGC	2526
Db	1917	TATAATCTCGCGCAACTATTTTCCCTCGAACACTTTTAAAGCCGTAGATAAACAGGCT	1976
Qy	2527	GTGGGTGACAGAGTACCATCACTGTAGGCCAGTTCAGGATGTGGTACTTCTGTAGCT	2586
Db	1977	GGGACACTTCAATGAGCGAAATAACATCGTCACCTGGGACATGTGTGACATCATGC	2036
Qy	2587	TGTCACCGACAGAAGCCAGGTAAAGCTCCAAAGCTGCTGATCTACTGGACATCCACCCGG	2646
Db	2037	ACGTAAACTCGCAAGCCGACTGATGC-----CTTCTGACAAATGGAAGGCAATTATTG	2089
Qy	2647	CACACTGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGATTTCACTTCAACC	2706
Db	2090	CCGTAAAGCGGTGCGGCTCTGGTACCGGTGGTGAAGACCAAGAACAGCACCTCGATC---	2146
Qy	2707	ATCAGAGCGCTCCAGCAGGACATCGCACTACTACTGCCAGCAATATAGCTCTAT	2766
Db	2147	-TGAGCCGGGATATTGCCCAGCGGTTTCAACGCGCTGTATGGCGAGATCGATCCCGTC---	2202
Qy	2767	CGGTCTTTCGGCCAGGGACCAAGGTGGAAATCAAAACGAGGTGGCTCAGGATCGGGTGA	2826
Db	2203	-----GTTTTACAGCTCGTGNCTGGGAACCCCTGGCGTTA	2239
Qy	2827	TCCGGCTGTGGGTCTCAGATCGGAGGTCCTAAGTGTGGAGAGCGGTGGAGGTGTGTG	2886
Db	2240	CCCAACTTAATGGCTTTGGAGACATCCCTTTCCGACGCTGGCGCTTAATAGCGAAGAGG	2299
Qy	2887	CAACCTGGCGGTCTCGCTCGCTCTCTCTCGCTCTGGCTTCCATTTCCACCATAT	2946
Db	2300	CCGCGACCGATCGCCCTTCCCAACAGTTGCGGACGCTGAAATGGCGGAAATGGCGCTTTGCC	2359
Qy	2947	TGGATGAGTTGGGTGAGACAGGCACCTGGAAAGGTCCTTGGTGGATTTGGAGAAATTCAT	3006
Db			
Db	2360	TGTTTTCCGGCACCAAGAGCGGTCCGGAAGCTGGCTGGAGTGCAT-----	2407
Qy	3007	CCAGATACAGTACGATTAACTATATGCGCCGCTCTCTAAAGGATAGATTTACAAATATCGCA	3066
Db	2408	--CTTCTTGAGGCCATATCTGTCTGCTCCCTCAACTGGCAGATGACGGTTAGATG	2465
Qy	3067	GACAAACGCAAGAAACACATTTGCTGCAAAATGGACAGCCTGAGACCCGAAAGACACCGGG	3126
Db	2466	CGCCCATCTACACCAACGCTGACCTATCCATTAAGGTCAATCCGCGCTTTGTTCCACGG	2525
Qy	3127	GTCTAATTTTGTGCAAGCCTTACTTTCGGCTTCCCTGGTTCTTATTTGGGGCCCAAGGG	3186
Db	2526	AGAAATCCGACGGGTTGTACTCGCTCACATTTTAATGTGTGAAGCTGGCTACAGGAA	2585
Qy	3187	ACCCCGGTCAACGCTCTCCAGTCTAAGCCCAACAGCAGCCAGCGCGCGACCAACA	3246
Db	2586	GGCCAGACCGCAATTAATTTTGTATGGCGTTAACTCGGCGTTTCACTCTG-----	2633
Qy	3247	CCGGCGCCCAACCATCGCGTCGAGCCCTGTCTCCCTGCGCCAGAGCGCGCTCGGCCAGCG	3306
Db	2634	-----TGGTGCAACGGGCGCTGGGTGGTTACGGGCAAGACAGTCTG	2674
Qy	3307	GGGGGGGCGAGTGCACACAGAGGGGCTGGACTTTCGCCCTGGATCCCAACTCTGTCTAC	3366
Db	2675	TTTGGCGTCTTAAATTTGAGCTCGAGCGCATATCTACGCGCCGGA-----GAAACCCGCTC	2730
Qy	3367	CTGCTGGATGGAATCCTCTTCACTATATGTGTCTTCTCACTGCTTGTCTCTGAGAGTG	3426
Db	2731	GGGTGTATGGTCTCGCTGGAGTGACGGGAGTTATCT-----	2768
Qy	3427	AAAGTTACAGAGAGCGAGAGCCCGCGTACACAGAGGGCCAGAACACAGCTCTATAAC	3486
Db	2769	-----TGAGATCAAGATATGTGGCGGATGAGCGGATTCAGAGCGAAAC	2814
Qy	3487	GAGCTCAATCTAGACGAAAGAGAGAGTACGATGTTTTGGACAAGAGAGCTGCCGGGAC	3546
Db	2815	GGTCTGCGCTGCGGACGCGGAAATGAATATATGGCCCAACACAGAGTGGGCGCGCGAC	2874
Qy	3547	CCTGAGATGGGGGAAGCCGAGAGGAAGAACCTCAGGAAGGCTGTACAAATGAACITG	3606
Db	2875	TTCCAGTT-----CAACATCAGCGCTACAGTCAACAGCAACTGATGGAACCGACCAT	2928
Qy	3607	CAGAAAGATAGATGCGGAGCGCTACAGTGTAGATTGGGATGAAAGCGCGCGCGAGG	3666
Db	2929	CGCCATCTCTGCAACGCGGAAGAACCGACATGCTGTTATACGACGTTTCCATATGGGG	2988
Qy	3667	GGCAAGGGGACAGATGGCTTTTACAGGGTCTCAGTACGACCACCAAGAGACACTACGAC	3726
Db	2989	ATTGGTGGCGACACTCTCTGGAGCCCGTCAAGTATCGGCGGAATTCAGCTGAGCGCGGT	3048
Qy	3727	GCCCTTCACTCAGCGCCCTGCCCCCTCGCTAACTCGACGCGCGCGGATCCGGATTAG	3786
Db	3049	CGCTACCATTTACAGTTGGTCTGGTGTCAAAATTAATAAATACCGGGGAGGCCATGCTG	3108
Qy	3787	TCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACAATATC	3846
Db	3109	CCCGTATTTCCGCTAAGGAAATCCATTATGT-----ACTATTTA	3147
Qy	3847	ACCAGCTGAAGCTATAGAGTACGAGCCATAGATAAAATAAAAGATTTTATTTAGTCTCC	3906
Db	3148	AACTCGAGCGCGCGCAGCACAGTGGTTCGACGATAAAATAAAAGATTTTATTTAGTCTCC	3207
Qy	3907	AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTTCGCAAGCTAGCTTAAGTAACGC	3966
Db	3208	AGAAAAGGGGGAATGAAGACCCACCTGTAGTTTTCGCAAGCTAGCTTAAGTAACGC	3267
Qy	3967	CAATTTGCAAGCATGG-AAAAATACATAACTGAGAATAGAGAAGTTTTCAGATCAAGTTAG	4025
Db	3268	CAATTTGCAAGCATGGAAAATACATAACTGAGAATAGAGAAGTTTTCAGATCAAGTTAG	3327
Qy	4026	GAAACAGA-GAGACAGAGAAATATGGGCCAAAACAGGATATCTGTGTGAAGCAGTTCTGCC	4084
Db	3328	GAAACAGTGAACAGCTGAATATATGGGCCAAAACAGGATATCTGTGTGAAGCAGTTCTGCC	3387

QY 4085 CC-GCTCAGGGCCAAAGACAGTTGGAAACAGAGGAATATATGGGCCAAACAGGATATCTGTGG 4143
Db |||||
QY 3388 CCGGCTCAGGGCCAAAGAACAGATGGAACAGCTGGAATATGGGCCAAACAGGATATCTGTGG 3447
Db |||||
QY 4144 TAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGTCCTCCAGATGGTCCCGCC 4203
Db |||||
QY 3448 TAAGCAGTTCTGCCCCGGCTCAGGGCCAAAGAACAGATGGTCCCGATGGTCCAGCC 3507
Db |||||
QY 4204 CTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCCAAGGACCTGGAATGAC 4263
Db |||||
QY 3508 CTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCCAAGGACCTGGAATGAC 3567
Db |||||
QY 4264 CCGTGCCTTATTTGAACCTAACCAATCAGTTGCTGCTTCTCGCTTCTGTTCCGGCGCTTCTG 4323
Db |||||
QY 3568 CCGTGCCTTATTTGAACCTAACCAATCAGTTGCTGCTTCTCGCTTCTGTTCCGGCGCTTCTG 3627
Db |||||
QY 4324 CTCCCGAGCTCAATAAAGAGCCACAAACCCCTCAGTCGGCGCCAGTCTCCGATAG 4383
Db |||||
QY 3628 CTCCCGAGCTCAATAAAGAGCCACAAACCCCTCAGTCGGGGGCCAGTCTCCGATG 3687
Db |||||
QY 4384 ACTGCGTCCCGGGTACCCGCTGTTCTCAATAAACCCCTCTTCAGATTCGATCCGACTCGT 4443
Db |||||
QY 3688 ACTGAGTCGCCGGGTACCCGCTGATCAATAAACCCCTCTTCGAGTTGCATCCGACTGT 3747
Db |||||
QY 4444 GGTCTGCTGTTCTTGGGAGGGTCT-CTCTGAGTGATGACTACCCGTCAGCGGGTCT 4502
Db |||||
QY 3748 GGTCTGCTGTTCTTGGGAGGGTCTCTCTGAGTGATGACTACCCGTCAGCGGG-- 3804
Db |||||
QY 4503 TTCAGTTCTCCACCTACAGGTCTCACTAACTTCTCTGATGTCGGCAGGACTCCG 4562
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4563 TCAGCCCGGTTTTTGTATATAAATGAAGAACAGTGTCTTCAAGCCAGACTAC 4622
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4623 ATCTGACTCTGGCTTTATAAAGATGTTGAAGGGCTGTGGACTATCTGCCACAG 4682
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4683 ACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCTGCTATT 4742
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4743 TTGCTGGCTTCTCGGTATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCAGGGA 4802
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4803 TTGAGAGTACTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGGTTGCTGGGCTCAGTG 4862
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4863 TGAGATTTTGCCATAAAGGGTCTGCCCCCTGTGTACAGACAGATCGGAATCTAGAT 4922
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4923 GCATACTCAGAGTCCCGCGTTCCGGGCTCTGTATCTCAGGGCATCTTTGCTAGAGAT 4982
Db |||||
QY 3805 ----- 3804
Db |||||
QY 4983 CCTCTACGCGGACGCATCTGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTGTCATAG 5042
Db |||||
QY 3805 ----- 3804
Db |||||
QY 5043 CTGTTTCTGTGAAATTTTATCCGCTCACAATTCACAACATACGAGCGGAGC 5102
Db |||||
QY 3805 ----- 3804
Db |||||
QY 5103 ATAAAGTGAAGCTTGGGTGCCTAATGAGTGAGCTAACTCACATTAAATGCGTGGC 5162
Db |||||
QY 3805 ----- 3804
Db |||||

QY 5163 TCACCTGCCCGCTTTCCAGTCCGGGAAACCTGTGCTGTCAGCTGATTAATGAAATCGGCCAA 5222
Db |||||
QY 3805 ----- 3804
Db |||||
QY 5223 CGCGCGGGAGAGCGGTTTCGGGTATTTGGGGCGCTCTTCGGCTTCCCTCCTCAGCTACGCTCG 5282
Db |||||
QY 3836 CGCGCGGGAGAGCGGTTTCGGGTATTTGGGGCGCTCTTCGGCTTCCCTCCTCAGCTACGCTCG 3895
Db |||||
QY 5283 CTGCGCTCGGTCTGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGCG 5342
Db |||||
QY 3896 CTGCGCTCGGTCTGCTGCGGCGAGCGGTATCAGCTCACTCAAAAGCGGTAAATAGCG 3955
Db |||||
QY 5343 TTATCCACAGAAATCAGGGGATAACGCGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG 5402
Db |||||
QY 3956 TTATCCACAGAAATCAGGGGATAACGCGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAG 4015
Db |||||
QY 5403 GCCAGGAACCGTAAAGAGCGGCTGCTGCGGCTTTTCCATAGGCTCCGCCCCCTGAC 5462
Db |||||
QY 4016 GCCAGGAACCGTAAAGAGCGGCTGCTGCGGCTTTTCCATAGGCTCCGCCCCCTGAC 4075
Db |||||
QY 5463 GAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGA 5522
Db |||||
QY 4076 GAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGA 4135
Db |||||
QY 5523 TACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTT 5582
Db |||||
QY 4136 TACAGCGGTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTGCGCTT 4195
Db |||||
QY 5583 ACCGGATACCTGTCGGCTTCTCCCTCGGGAGCGTGGCGCTTCTCATAGCTCAGC 5642
Db |||||
QY 4196 ACCGGATACCTGTCGGCTTCTCCCTCGGGAGCGTGGCGCTTCTCATAGCTCAGC 4255
Db |||||
QY 5643 TGTAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCCAAGCTGGGCTGTGTGACGAAACC 5702
Db |||||
QY 4256 TGTAGGTATCTCAGTTCGGTGTAGGTGTTCCGCTCCAAGCTGGGCTGTGTGACGAAACC 4315
Db |||||
QY 5703 CCCCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCAACCCCGTA 5762
Db |||||
QY 4316 CCCCTTCAGCGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCAACCCCGTA 4375
Db |||||
QY 5763 AGACACACTTATCGCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGGTAT 5822
Db |||||
QY 4376 AGACACACTTATCGCACTGGCAGCAGCCACTGGTAAACAGGATTTAGCAGAGCGAGGTAT 4435
Db |||||
QY 5823 GTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACCGCTACCTAGAGGACA 5882
Db |||||
QY 4436 GTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTAACTACCGCTACCTAGAGGACA 4495
Db |||||
QY 5883 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCT 5942
Db |||||
QY 4496 GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCT 4555
Db |||||
QY 5943 TGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGACAGCAGATTT 6002
Db |||||
QY 4556 TGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGACAGCAGATTT 4615
Db |||||
QY 6003 ACAGCGAGAAAAAAGGATCTCAAGAGAGTCTTGTGATCTTTCTACGGGCTCTGACGCT 6062
Db |||||
QY 4616 ACAGCGAGAAAAAAGGATCTCAAGAGAGTCTTGTGATCTTTCTACGGGCTCTGACGCT 4675
Db |||||
QY 6063 CAGTGGAAACGAAACTCACCGTTAAGGATTTGGTCTAGAGATTTATCAAAAGGATCTTC 6122
Db |||||
QY 4676 CAGTGGAAACGAAACTCACCGTTAAGGATTTGGTCTAGAGATTTATCAAAAGGATCTTC 4735
Db |||||
QY 6123 ACCTAGATCTTTTAAATTTAAATGAAGTTTAAATCAATCAATATATATATAGTAA 6182
Db |||||
QY 4736 ACCTAGATCTTTT- ----- GCGCGCGCGCGCAAAATCAATCTAAAGTATATATAGTAA 4788
Db |||||
QY 6183 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTCTA 6242
Db |||||
QY 4789 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTCTA 4848
Db |||||
QY 6243 TTTGCTTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGAGGCG 6302
Db |||||

Db 4849 TTTTCGTTTCCATAGTTCCTGACTCCCGCTCGTGTAGATAACTACGATACGGAGGCG 4908
Qy 6303 TTACCATCTGGCCCGAGTCTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGAT 6362
Db 4909 TTACCATCTGGCCCGAGTCTGCAATGATACCGGAGACCCAGCTCACGGCTCCAGAT 4968
Qy 6363 TTATCAGCAATAAACCAGCAGCCGGAAGGGCGGAGCAGAGTGGTCTGCAACTTTA 6422
Db 4969 TTATCAGCAATAAACCAGCAGCCGGAAGGGCGGAGCAGAGTGGTCTGCAACTTTA 5028
Qy 6423 TCGGCTTCCATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAGTTCGCGAGTT 6482
Db 5029 TCGGCTTCCATCCAGTCTATTAAATGTTGCGGGAAGCTAGAGTAAGTTCGCGAGTT 5088
Qy 6483 AATAGTTTGGCAACCTGTTGCGCATTTGTACAGGC-TGTTGGTGTACGCTCGTCTTT 6541
Db 5089 AATAGTTTGGCAACCTGTTGCGCATTTGTACAGGCATCGTGTGTACGCTCGTCTTT 5148
Qy 6542 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATG 6601
Db 5149 GGTATGGCTTCATTTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCGCATG 5208
Qy 6602 TTTTGCAAAAAGCGGTAGTCTCTTCGCTCCGATCGTTGTTCAGAAAGTAAAGTTGGCC 6661
Db 5209 TTTTGCAAAAAGCGGTAGTCTCTTCGCTCCGATCGTTGTTCAGAAAGTAAAGTTGGCC 5268
Qy 6662 GCAGTGTATCACTCATGTTATGGCAGCAGTGCATATAATCTCTTACTGTCTATGCCATCC 6721
Db 5269 GCAGTGTATCACTCATGTTATGGCAGCAGTGCATATAATCTCTTACTGTCTATGCCATCC 5328
Qy 6722 GTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCACTCTGAGAATAGTGTAG 6781
Db 5329 GTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCACTCTGAGAATAGTGTAG 5388
Qy 6782 CGCGCAGCGAGTTGCTTCTGCGCGGTCAATACGGGATAATACCGCGCACATAGCAGA 6841
Db 5389 CGCGCAGCGAGTTGCTTCTGCGCGGTCAATACGGGATAATACCGCGCACATAGCAGA 5448
Qy 6842 ACTTTAAAGTGTCTATCATTTGGAACCGTTCTTGGGGCGGAAAACCTCTCAAGGATCTTA 6901
Db 5449 ACTTTAAAGTGTCTATCATTTGGAACCGTTCTTGGGGCGGAAAACCTCTCAAGGATCTTA 5508
Qy 6902 CGCGTGTGAGTCCAGTTTCGATTAACCCACTCGTGCACCACTGATCTTCAGCATCT 6961
Db 5509 CGCGTGTGAGTCCAGTTTCGATTAACCCACTCGTGCACCACTGATCTTCAGCATCT 5568
Qy 6962 TTTTACTTTCCACGACGCTTCTGCGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAG 7021
Db 5569 TTTTACTTTCCACGACGCTTCTGCGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAAG 5628
Qy 7022 GGAATAAGGGCGACACGGAATTTGAATACATCTCTTCTTTTCAATATTATTTGA 7081
Db 5629 GGAATAAGGGCGACACGGAATTTGAATACATCTCTTCTTTTCAATATTATTTGA 5688
Qy 7082 AGCATTTATCAGGGTTATTTGTTCTCATGCGGATACATATTGAAATGATTTTGAATAAT 7141
Db 5689 AGCATTTATCAGGGTTATTTGTTCTCATGCGGATACATATTGAAATGATTTTGAATAAT 5748
Qy 7142 AAACAAATAGGGGTTCCGCGCACATTTCC 7170
Db 5749 AAACAAATAGGGGTTCCGCGCACATTTCC 5777

RESULT 8
ID AAX33182
XX AAX33182 standard; DNA; 7372 BP.
AC AAX33182;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid pRx-Bcl-xl-bsr.

XX Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
KW bcl-2; bcl-xl; FlIP; survivin; IAP; ILP; adenovirus; cancer;
KW autoimmune disease; graft rejection reaction; inflammation;
KW inflammatory disease; ss.
XX Synthetic.
OS Homo sapiens.
XX WO9913073-A2.
XX 18-MAR-1999.
XX 07-SEP-1998; 98WO-JP004010.
XX 08-SEP-1997; 97JP-00259235.
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX Hamada H;
XX WPI; 1999-243728/20.
XX New apoptosis-resistant virus-sensitive cell.
XX Example 2; Page 41-45; 51pp; English.
XX The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRx-Bcl-xl-bsr, which contains the human Bcl-xl gene, and is used in an example from the present invention
SQ Sequence 7372 BP; 2353 A; 1749 C; 1649 G; 1621 T; 0 U; 0 Other;
Query Match 41.6%; Score 3183.2; DB 2; Length 7372;
Best Local Similarity 70.7%; Pred. No. 0;
Matches 4893; Conservative 0; Mismatches 1503; Indels 528; Gaps 28;
Qy 843 TTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAATGACCTTGCT 902
Db 467 TGCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAATGACCTTGCT 526
Qy 903 TATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTCGCGCTTCTGCTCCCGAG 962
Db 527 TATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTCGCGCTTCTGCTCCCGAG 586
Qy 963 CTCATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCCGATAGACTCGCTCG 1022
Db 587 CTCATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCCGATAGACTCGCTCG 646
Qy 1023 CCCGGTACCGGTATTTCCCAATAAGCCTCTTGCTGTTTGCATCCGAATCGGACTCG 1082
Db 647 CCCGGTACCGGTATTTCCCAATAAGCCTCTTGCTGTTTGCATCCGAATCGGACTCG 705
Qy 1083 TGATCTTTGGGAGGCTCTCTCAGATTGATTGACTGCC-ACCTCGGGGTCTTTCATTT 1141

706	TGTTTCCTTGGGAGGGTCTCTCTGAGTGATGTGACTACCGGTACGCGGGGTCTTTTCATTT	765
1142	GGAGGTTCCACCGAGATTTTGGAGACCCCTGCTCCAGGGAGCACCGACACCCGCCCGCGGAG	1201
766	GGGGCTCGTCCGGGATCGGGAGACCCCTGCTCCAGGGAGCACCGA-CCCACCAACCGGAG	824
1202	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTTGTCTAGTGTCTATGACTGATTTT	1261
825	GTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTTGTCTAGTGTCTATGACTGATTTT	884
1262	ATGCGCTCGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTTGGCGGACCCGTGGTGGGA	1321
885	ATGCGCTCGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTTGGCGGACCCGTGGTGGGA	944
1322	ACTGACGAGTTTCGAAACACCCGGCCGCAACCTCTGGGAGACGTCCCAGGGACTTTCGGGGCC	1381
945	ACTGACGAGTTTCGAAACACCCGGCCGCAACCTCTGGGAGACGTCCCAGGGACTTTCGGGGCC	1004
1382	CGTTTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTTAGGACTCTTTGGTGGCAAC	1441
1005	CGTTTTTGTGGCCGACCTGAGTCTCTAAATCCCGATCGTTTTAGGACTCTTTGGTGGCAAC	1064
1442	CCCCTTAGAGGAGGATATGTGGTTCTGTGTAGGAGACGAGAACTCTAAAAAGTTTCCGCC	1501
1065	CCCCTTAGAGGAGGATATGTGGTTCTGTGTAGGAGACGAGAACTCTAAAAAGTTTCCGCC	1124
1502	TCCGCTCGAATTTTGTCTTTCCGTTTCGGGACCGAAGCCGCGCGCGGTCTTGTCTGCTG	1561
1125	TCCGCTCGAATTTTGTCTTTCCGTTTCGGGACCGAAGCCGCGCGGTCTTGTCTGCTG	1184
1562	CAGCATCGTTCTGTGTGTCTCTGTCTGTGACTGTGTCTGTATTTGTCTGAAAATATGGG	1621
1185	CAGCATCGTTCTGTGTGTCTCTGTCTGTGACTGTGTCTGTATTTGTCTGAAAATATAGGG	1244
1622	CCCGGGCTAGACTGTTACCACTCCCTTAAAGTTTGAACCTTAGGTCACCTGGAAAGATGTCGA	1681
1245	CC-----AGACTGTATACCACTCCCTTAAAGTTTGAACCTTAGGTCACCTGGAAAGATGTCGA	1298
1682	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAGAGACGTTGGGTGTTACCTTCTGCTC	1741
1299	GCGGATCGCTCACAAACAGTCGGTAGATGTCAAGAGAGACGTTGGGTGTTACCTTCTGCTC	1358
1742	TGCAGAAATGCCAAACCTTTAAACGTCGATGGCCGCGAGACGGCACTTTTAAACCGAGACCT	1801
1359	TGCAGAAATGCCAAACCTTTTAAACGTCGATGGCCGCGAGACGGCACTTTTAAACCGAGACCT	1418
1802	CATCACCCAGGTTAAGATCAAGTCTTTTACCTGGCCCGGATGGACACCCAGACCAAGT	1861
1419	CATCACCCAGGTTAAGATCAAGTCTTTTACCTGGCCCGGATGGACACCCAGACCAAGT	1478
1862	CCCCTACATCGTGACCTGGGAAGCTTTGGCTTTTGGACCCCTCCCTGGGTCAAGCCCTT	1921
1479	CCCCTACATCGTGACCTGGGAAGCTTTGGCTTTTGGACCCCTCCCTGGGTCAAGCCCTT	1538
1922	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCCATCGCCGCCGTCTCTCCCTCTGAAACC	1981
1539	TGTACACCTTAAGCCTCGGCTCTCTTCTCTCCATCGCCGCCGTCTCTCCCTCTGAAACC	1598
1982	TCCTCGTTGACACCCGGCTTGATCTCCCTTTATTCGAGCCCTCATCTCTTCTCTAGGCGC	2041
1599	TCCTCGTTGACACCCGGCTTGATCTCCCTTTATTCGAGCCCTCATCTCTTCTCTAGGCGC	1658
2042	CCCCTATGGCCATATGAGATCTTATATGGGGACCCCGCCCTTGTAACTTCCCTGA	2101
1659	-----CCATATGAGATCTTATATGGGGACCCCGCCCTTGTAACTTCCCTGA	1708
2102	CCCTGCATGACAAGAGTTACTTAACAGCCCTCTCTCTCAAAGCTCACCTTACAGGCT-TCCTA	2160
1709	CCCTGCATGACAAGAGTTACTTAACAGCCCTCTCTCTCAAAGCTCACCTTACAGGCTCTCTA	1768
2161	CTTAGTCCAGCAAGTCTGGAGACTCTTGGCGGACGCTTACAGAAACAACTGGACCG	2220
1769	CTTAGTCCAGCAAGTCTTGGAGACTCTTGGCGGACGCTTACAGAAACAACTGGACCG	1828

Qy	2221	ACCGGTGGTACTCTCA	CCCTTTACCGAGTCGGCGACACAGTGTGGTGGCGGACACAGAC	2228
Db	1829	ACCGGTGGTACTCTCA	CCCTTTACCGAGTCGGCGGACACAGTGTGGTGGCGGACACAGAC	1888
Qy	2281	TAGAAGCTAGAACCTCGCTGGAAAGACCTTACACAGCTCTCTGACCAACCCCAACCGC	2340	
Db	1889	TAGAAGCTAGAACCTCGCTGGAAAGACCTTACACAGCTCTCTGACCAACCCCAACCGC	1948	
Qy	2341	CCTCAAGTAGACGGCATCGACGCTTGGATACACGCGCCGCCACGTAAGGCTGCCGACCC	2400	
Db	1949	CCTCAAGTAGACGGCATCGACGCTTGGATACACGCGCCGCCACGTAAGGCTGCCGACCC	2008	
Qy	2401	CGGGGGTGGACCAATCTCTTAGACCTGCCATGGGATGAGCTGTATCATCTCTTCTTCTTGGTGA	2460	
Db	2009	CGGGGGTGGACCAATCTCTTAGACCTGCCATGGGATGAGCTGTATCATCTCTTCTTCTTGGTGA	2068	
Qy	2461	GCACACGCTACAG-----	2473	
Db	2069	AAAAATGTCTCAGAGCAACCGGAGCTGGTGGTGTGACTTTCTCTCTCAAGCTTTCCCA	2128	
Qy	2474	-----	2473	
Db	2129	GAAGGATACAGCTGGGATCAGTTTAGTGATGTGGAGAGAACAGGACTGAGGCCCCAGA	2188	
Qy	2474	-----	2480	
Db	2189	AGGACTGAATCGGAGATGGAGACCCCACTGCATCAATGGCAAACCATCTCTGGCACT	2248	
Qy	2481	CTCGGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGAGGCCAGCGTGGGTGACAGAGT	2540	
Db	2249	GGCAGACAGCCCGCGGTGAATGGAGCCACTGCGCACAGCAGCAGGTTTGGATGCCGGGA	2308	
Qy	2541	GACCATCACCTCTAAGGC-----	2577	
Db	2309	GGTGATCCCCATGGCAGCAGTAAGCAGCGCTGAGGAGGCAAGCGACGAGTTTGAAT	2368	
Qy	2578	TCTGTAGCTTGGTACACGACGAAGCCAGGTA---	2634	
Db	2369	CGGTCACGGCGGCATTTCACTGACCTGACATCCAGCTCCACATCACCCAGGGACAGC	2428	
Qy	2635	ACATCCACCCGGCACA	CTGTGTGCCAGAGATTCAGCGGTACGGTA	2687
Db	2429	ATATCAGAGCTTTGAACAGGCTAGTGAATGAATCTCTTCCGGGATGGGTAACTGGGGTCG	2488	
Qy	2688	TACCGACTTACCTTCCACATCAGCAGCCTCCAGCCAGCAGGACATGCCACCTACTACTG	2747	
Db	2489	CATTGTGGCCTTTTTCTCTTTCGGCGGGGCACTGTGCGTGGAAAGCGTAGACAAGAGAT	2548	
Qy	2748	CCAGCAATA-----	2756	
Db	2549	GCAGGTATTGGTGAGTCGGATCGCAGCTTGGATGGCCACTTACTGATGACCACTAGA	2608	
Qy	2757	---TAGCCTCTATCGGTCTGTCGGCCAAAGGGA	CCAAAGGTGGAATCAACAGAGTGGCTC	2813
Db	2609	GCCTTGGATCCAGGAGAACGGCGGCTGGGATCTTTTGTGGAACCTCTATGGGAACAATGC	2668	
Qy	2814	AGGATC---GGGTGGATCGGCTCTGGTGGCTCAGGATCGGAGGTCCTACTGCTGGAGAG	2870	
Db	2669	AGCAGCCGAGAGCCGAAGGGCCAGGAACGCTTCAACCGCTGTTTCTGACGGGCATGAC	2728	
Qy	2871	CGGTGGAGGTGTTGTGCAACCTGCGCGTCCCTGCGCCTG-----	2910	
Db	2729	TGTGGCCGGTGGTCTCTGTGGGCTCACTCTTTCAGTCGGAATGAAAAAAGAAAT	2788	
Qy	2911	TCCTGTCTCCGATCTGGCTTTTCGATTTTCCACACATATTTGGATGAGTTGGGTGAGACAGGCA	2970	
Db	2789	TCGAGCTCGGTACCCGGGGCGCGCGATCTCTTAATTAAGCGCGCGGTAAACAATGT	2848	
Qy	2971	CCTGGAAAAAGTCTTGATGGATTTGGAGAAATTCATCCAGATAGTAGTACGATTAATCTAT	3030	
Db	2849	TAACTACTAACTAAGCTAGCAACGGTTTTCCCTCTAGCGGATCAATTTCCCTCTCCCT	2908	

Qy	7173	GAAAAGTGCCACCTGAGCGTCTAGAAACCAATTATTCATGACATTAACTTATATAAATA	7232
Db	7289	GAAAAGGCCACCTGAGCGTCTAGAAACCAATTATTCATGACATTAACTTATATAAATA	7348
Qy	7233	GGCGTATCAGAGGCCCTTTTCGTC	7256
Db	7349	GGCGTATCAGAGGCCCTTTTCGTC	7372

RESULT 9	
AAAL57215	AAAL57215 standard; DNA; 6253 BP.
ID	AAAL57215 standard; DNA; 6253 BP.
XX	
XX	
AC	AAAL57215;
XX	
XX	25-SEP-2003 (first entry)
DT	
XX	
DE	MIG retroviral vector DNA sequence.
XX	
XX	MIG retroviral vector; lymphocyte; cytostatic; antiallergic;
KW	immunosuppressive; antimicrobial; anti-HIV; virucidal; gene therapy;
KW	cancer; allergy; autoimmune disorder; infectious disease; HIV; AIDS;
KW	OTII T cell receptor; TCR; ds.
XX	
XX	Unidentified.
OS	
XX	WO2003050262-A2.
PN	
XX	
XX	19-JUN-2003.
PD	
XX	
XX	10-DEC-2002; 2002WO-US039527.
PF	
XX	
XX	10-DEC-2001; 2001US-0339375P.
PR	
PR	08-JUL-2002; 2002US-0394803P.
XX	
XX	(CALY) CALIFORNIA INST OF TECHNOLOGY.
PA	
XX	
PI	Yang L, Van Parijs L, Baltimore D;
XX	
XX	WPI; 2003-532909/50.
DR	

612	Qy	GGAAATACATTAACCTGAGAAATAGAGAAGTTCCAGATCAACAGGTTAGGAAACAGAGACAGCA	671
61	Db	GGAAATACATTAACCTGAGAAATAGAGAAGTTCCAGATCAACAGGTTAGGAAACAGAGACAGC	119
672	Qy	GAATATGGGCCAAACAGGATATCTGTGGTGAAGCAGTTCTCTGCCCGCTCAGGSCCAAGAA	731
120	Db	-----	119
732	Qy	CAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATCTGTGTGTAAGCAGTTCTCTGCCCG	791
120	Db	-----AGAAATATGGSCCAAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCG	166
792	Qy	GCTCAGGGCCNAGAAACAGATGTCCTCCAGATGGGTCCGCCCTCAGCAGTTCTTAGAGA	851
167	Db	GCTCAGGGCCNAGAAACAGATGTFGCCAGATGGGTCCGCCCTCAGCAGTTCTTAGAGA	226
852	Qy	ACCATCAGATGTTTCCAGGGTGCCCAAGGACCTG-AAATGACCCCTGTGCTTATTTGAA	910
227	Db	ACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGCTTATTTGAA	286
911	Qy	CTAAACCAATCAGTTTCGCTTCTGTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA	970
287	Db	CTAAACCAATCAGTTTCGCTTCTGTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA	346
971	Qy	AGAGGCCCAACACCCCTCACTCGGCGCGCAGTCCTCCGATAGACGTGCGCCCGGTA	1030
347	Db	AGAGGCCCAACACCCCTCACTCGGCGCGCAGTCTCCGATAGACGTGCGTCGCC--GTA	405
1031	Qy	CCGGTATTCCCAATAAAGCCTCTGTGTTGTCATCCGAATCGTGAATCGCTGATCCTT	1090
406	Db	CCGGTATTCCCAATAAAGCCTCTGTGTTGTCATCCGAATCGTGAATCGCTGATCCTT	465
1091	Qy	GGGAGGCTCTCTCAGATTGATGACTGCCCCACTCGGGGTCTTTCAATTGGAGGTTCC	1150
466	Db	GGGAGGCTCTCTCAGATTGATGACTGCCCCACTCGGGGTCTTTCAATTGGAGGTTCC	525
1151	Qy	ACCGAGATTGGAGACCCCTGCCCCAGGACCAACCGACCCCCCGCGGGAGGTAAAGCTGG	1210
526	Db	ACCGAGATTGGAGACCCCTGCTAGGGACCAACCGACCCCCCGCGGGAGGTAAAGCTGG	585
1211	Qy	CGAGCAACTATCTGTGTCGTGTC-----CGATTGTCGTAGTGTCTATGACTGATTTTA	1262
586	Db	CGAGCGGTGTTTCGTGTCTGTCTCTGTCTTTGTGGGTGTTGTGCGCGCATCTAATGTT	645
1263	Qy	TGGCGCTCGGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGAA	1322
646	Db	TGGCGCTCGGTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGAA	705
1323	Qy	CTGACGAGTTCGGAAACACCCGGCCGCAACCTTGGGAGACGTCTCCAGGACCTTCGGGGGCC	1382
706	Db	CTGACGAGTTCGGAAACACCCGGCCGCAACCTTGGGAGACGTCTCCAGGGAATTTTGGGGGCC	765
1383	Qy	GTTTTTTGGCCCGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTGGTGCACCC	1442
766	Db	GTTTTTTGGCCCGACCTGAGGAAGGGAGTCGATGTGGAATCCGACCCCGTCA-----	818
1443	Qy	CCCTTAGAGAGGAGATATGTGGTCTTGTGTAGGAGACGAGAACCTTAAACAGTTCCCGCCT	1502
819	Db	-----GGATATGTGGTCTTGTGTAGGAGACGAGAACCTTAAACAGTTCCCGCCT	866
1503	Qy	CGGTCTGAATTTTGTCTTTCGGTTTGGGACCGAAGCCGCGCGCTCTTGT---CTGC	1559
867	Db	CGGTCTGAATTTTGTCTTTCGGTTTGGGACCGAAGCCGCGCGCTCTTGTCTGCTGCAAGCC	926
1560	Qy	TGCAGCATCGTTCGTGTGTTGTCCTCTCTGACTGTGTTTCTGTATTTGTCGAAAAATATG	1619
927	Db	TGCAGCATCGTTCGTGTGTTGTCCTCTCTGACTGTGTTTCTGTATTTGTCGAAAAATTAG	986
1620	Qy	GGCCCGGGCTAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTC	1679
987	Db	GGCC-----AGACTGTATACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTC	1040

PT equivalent.

XX Disclosure; SEQ ID NO 4; 108pp; English.

PS The invention relates to a novel method for enhancing the survival of

XX neurons and/or of cells expressing tyrosine hydroxylase (SC 1.14.16.2 - Tyrosine 3-monooxygenase) (TH +). The method comprises contacting a population of cells with Bcl-XL or its functional equivalent, where the population of cells is selected from: neurons or cells capable of differentiating into neurons; or TH expressing cells or cells capable of differentiating into TH expressing cells. The invention further comprises: a composition of cells obtainable by the method above; a composition of isolated mammalian cells overexpressing the anti-apoptotic Bcl-XL protein; a neural progenitor cell; a differentiated dopaminergic neuron; an implantable cell culture device comprising: a semi-permeable membrane permitting the diffusion of a biologically active protein through it; and a composition of cells selected from above; a lentiviral vector particle being produced based on a lentiviral transfer vector; enhancing the survival of TH + cells in vivo; a retroviral particle being produced based on a retroviral transfer vector; enhancing the survival of in vivo differentiated dopaminergic neurons; a packaging cell line capable of producing an infective vector particle; a packaging cell line capable of producing an infective vector particle; treatment of a neurological disorder; a fusion protein comprising the Bcl-XL sequence comprising 233 amino acids ADQ80670 or its functional equivalent and a membrane translocation signal; an expression vector comprising a polynucleotide sequence coding for the fusion protein and a promoter sequence capable of directing the expression of the fusion protein in a host cell; a host cell comprising the expression vector; and producing the fusion protein. The compositions of the invention have neuroprotective, neurotropic, and antiparkinsonian activities. The cells are useful for transplantation, drug screening, gene profiling, or for the preparation of a medicament useful for the treatment of a CNS disorder. The CNS disorder is a neurodegenerative disease involving lesioned and traumatic neurons, including traumatic lesions of peripheral nerves, the medulla, the spinal cord, cerebral ischaemic neuronal damage, neuropathy, peripheral neuropathy, Alzheimer's disease, Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis, or memory impairment connected to dementia. The method is useful for enhancing the survival of neurons and/or of cells expressing tyrosine hydroxylase for the treatment of neurodegenerative disorders. This sequence represents an expression construct for directing the expression of Bcl-XL in transduced cells, used in the method for increasing the survival rate of neurons of the invention.

XX Sequence 7257 BP; 1712 A; 2023 C; 1892 G; 1630 T; 0 U; 0 Other;

SQ

Query Match 40.9%; Score 3130.2; DB 13; Length 7257;

Best Local Similarity 69.6%; Pred. No. 0;

Matches 4939; Conservative 0; Mismatches 1763; Indels 391; Gaps 36;

QY 550 TTTGAAAGACCCACCTGTAGGTTGGGAGCTAGCTTAAGTAAACGCCATTTTGCAGGC 609

DB 167 TGTGAAAGACCCACCTGTAGGTTGGGAGCTAGCTTAAGTAAACGCCATTTTGCAGGC 226

QY 610 ATGG-AAAATACATACTCAGATACAGAGTTTCAGATCAAGTTAGGAACAGA-GAGAC 667

DB 227 ATGGAATAATACATACTCAGATACAGAGTTTCAGATCAAGTTAGGAACAGA-GAGAC 286

QY 668 AGCAGAAATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTCTGCCCC-GCTCAGGGCC 726

DB 287 AGCTGAATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTCTGCCCCGCTCAGGGCC 346

QY 727 AAGAACAGTTGGAAACAGGAGATATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTCTG 786

DB 347 AAGAACAGATGGAAACAGCTGAATATATGGGCCAAACAGGATATCTGTGGTAAAGCAGTTCTCTG 406

QY 787 CCCGGCTCAGGGCCAGAACAGATGTCCTCCAGATGGCGGTCCCGCTCCAGCAGTTTCT 846

DB 407 CCCGGCTCAGGGCCAGAACAGATGTCCTCCAGATGGCGGTCCCGCTCCAGCAGTTTCT 466

QY 847 AGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGCCTTATT 906

DB 467 AGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGCCTTATT 526

QY 907 TGAATTAACCATCAGTTTCCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCA 966

DB 527 TGAATTAACCATCAGTTTCCGCTTCTCGCTTCTGTTGCGCGCTTCTGCTCCCGAGCTCA 586

QY 967 ATAAAGAGAGCCCAACACCCCTCAGTCGCGCGCCAGTCTCTCCGATAGACTGCGTGCCTCG 1026

DB 587 ATAAAGAGAGCCCAACACCCCTCAGTCGCGCGCCAGTCTCTCCGATAGACTGCGTGCCTCG 646

QY 1027 GGTACCCGTTATCCCAATAAAGCCTTCTGCTGTTTTCATTCGAAATCGTGGATCTCGCTGAT 1086

DB 647 GGTACCCGTTATCCCAATAAAGCCTTCTGCTGTTTTCATTCGAAATCGTGGATCTCGCTGAT 705

QY 1087 CTTTGGGAGGCTCTCTCAGATGTTGATGACTGCCCC-ACTCGGGGGTCTTTTCATTTGGAG 1145

DB 706 CTTTGGGAGGCTCTCTCAGATGTTGATGACTGCCCC-ACTCGGGGGTCTTTTCATTTGGGG 765

QY 1146 GTTCCACCGAGATTTGGAGACCCCTGCGCCAGGAGACCCAGACCCCGCCCGGGAGGTAA 1205

DB 766 GTTCCACCGAGATTTGGAGACCCCTGCGCCAGGAGACCCAGACCCCGGGAGGTAA 824

QY 1206 GCTGCCAGCAACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265

DB 825 GCTGCCAGCAACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884

QY 1266 GCTGCCAGCAACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1325

DB 885 GCTGCCAGCAACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944

QY 1326 ACAGTGTCCGAAACACCCCGCCAGGAGACCCCTGCGGAGAGCTCCCAAGGAGACTTCGCGGGCCGTT 1385

DB 945 ACAGTGTCCGAAACACCCCGCCAGGAGACCCCTGCGGAGAGCTCCCAAGGAGACTTCGCGGGCCGTT 1004

QY 1386 TTTTGGGCCGACCTGAGTCTTAAATCCGATCGTTTATAGGACTCTTTGGTGCACACCCCTCC 1445

DB 1005 TTTTGGGCCGACCTGAGTCTTAAATCCGATCGTTTATAGGACTCTTTGGTGCACACCCCTCC 1064

QY 1446 TTAGAGGAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1505

DB 1065 TTAGAGGAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1124

QY 1506 TCTGAATTTTCTGCTTTCGTTTGGGACCGAAGCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1565

DB 1125 TCTGAATTTTCTGCTTTCGTTTGGGACCGAAGCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184

QY 1566 ATCGTTTCTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1625

DB 1185 ATCGTTTCTGTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1244

QY 1626 GGCTAGACTGTTTACCACTCCCTTAAAGTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1685

DB 1245 GGCCAGACTGTTTACCACTCCCTTAAAGTTTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304

QY 1686 ATCGCTTCAACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1745

DB 1305 ATCGCTTCAACACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364

QY 1746 GAATCGCCAACTTTTAAAGTGGCGCGAGACCGGACCTTTTAAACCGAGACCTCTATC 1805

DB 1365 GAATCGCCAACTTTTAAAGTGGCGCGAGACCGGACCTTTTAAACCGAGACCTCTATC 1424

QY 1806 ACCGAGTTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1865

DB 1425 ACCGAGTTTAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484

QY 1866 TACATCTGACTCGGAAAGCCTTGGCTTTTGACCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1925

DB 1485 TACATCTGACTCGGAAAGCCTTGGCTTTTGACCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1544

QY 1926 CACCTTAAGCCTCGCCTCTCTTCTCCTCATCCGCGCGCTCTCTCCCTCCCTTGAACCTCTCT 1985

DB 1545 CACCTTAAGCCTCGCCTCTCTTCTCCTCATCCGCGCGCTCTCTCCCTCCCTTGAACCTCTCT 1604

QY 1986 CGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTTAGCGCCGCC 2045
DB 1605 CGTTCGACCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTTAGCGCCGCC 1664
QY 2046 ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGACCT 2105
DB 1665 ATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGACCT 1724
QY 2106 GACATGACAAGATTACTAACAAGCCCTCTCTCCAAAGCTCACTTACAGGCTCTCTACTTA 2164
DB 1725 GACATGACAAGATTACTAACAAGCCCTCTCTCCAAAGCTCACTTACAGGCTCTCTACTTA 1784
QY 2165 GTCCAGCAGGAGTCTGAGACCTCTGGGGCAGCTTACCAAGAACAACTGGACCGACG 2224
DB 1785 GTCCAGCAGGAGTCTGAGACCTCTGGGGCAGCTTACCAAGAACAACTGGACCGACG 1843
QY 2225 GTGTGACCTCACCTTACCGAGTCCGGGACACAGTGTGGGTCCGGCCGACACCACTAAAG 2284
DB 1844 GTGTGACCTCACCTTACCGAGTCCGGGACACAGTGTGGGTCCGGCCGACACCACTAAAG 1903
QY 2285 AACCTAGAACCTCGTGGAAAGACCTTACACAGTCTCTGCTGACACCCGCCCTC 2344
DB 1904 AACCTAGAACCTCGTGGAAAGACCTTACACAGTCTCTGCTGACACCCGCCCTC 1963
QY 2345 AAAGTAGAGGATCGAGCTTGGATACACGGCCGCCAGTGAAGGCTGCCGACCCCGGG 2404
DB 1964 AAAGTAGAGGATCGAGCTTGGATACACGGCCGCCAGTGAAGGCTGCCGACCCCGGG 2023
QY 2405 GGTGACCATCTCTAGACTGCGATGGGATGGAGTGTATCATCTCTTCTTGTAGCAA 2464
DB 2024 GGTGACCATCTCTAGACTGCGGATCGAATTTGGTACCGAGCTCGGATC-CACTAGTAA 2082
QY 2465 CAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTTGAGCGCA 2524
DB 2083 CGGCGCCAGTGTCTGAGGCCACTCTCTATTATAAATGTCTCAGAGCAACCGGAGC 2142
QY 2525 CGTGGGTGACAGAGTGAGC-ATCACTGTGAAGCCAGTCAAGATGTGGGTACTTCTGTA 2583
DB 2143 TGGTGGTGTACTTCTCTCTACAGCTCTCCAGAAAGGATACAGCTGGAGTCAGTTTA 2202
QY 2584 GCTTGGTACCAGAGAACCGAGGTAAAGCTCCAAAGCTGTGATCTACTGACATCCACC 2643
DB 2203 CGGATGTGGAAGAGAACAGACTGAAGCCCGAGAGAACTGAACAGAAAGGAGACCC 2262
QY 2644 CGGCACACTGTGTGCCAAGCAGATTACAGCGGTAGCGGTAGCGGTACCGACTTCACTTC 2703
DB 2263 CCAGTGCCATCAATGGCAACCATCTCTGGACCTGGCGGATAGCCCCGGGTGATGGAG 2322
QY 2704 ACCATCAGAGCTCCAGCAGAGACATCGCCACTACTACTGCGCAATATAGCCTC 2763
DB 2323 CCACTGGCCACAGCAGCAGTTTGGATGGCGCGGAGGTACTCCCATGGCAGCAGTGAAGC 2382
QY 2764 TATCGGTGTTGGCCAGGAGACCAAGT-----GGNAATCAACAGAGTGG 2810
DB 2383 AAGCGCTGAGAGAGCTGCGATGAGTTGAACTGCGGTACCGGAGCATTCAGTGATC 2442
QY 2811 CTCAGGATCGGTGATCGGCTCTGTGGTCTCAGGATCGGAG-GTCCAACTGTGGAGA 2869
DB 2443 TAAATCCAGCTTCATATAACCCAGGAGCAGTATATCAGAGCTTTGAAACAGGTATGTA 2502
QY 2870 CGGTGGAGGTGTTGTGCAACTGCGCGGTCCCTGCGCTGTCTCTGCTCCGCATCTGGCT 2929
DB 2503 ATGAACCTTTTCGGGATGGGTAAACTTGGGTGCGATTGTGGCTCTCTCTCTTTGGGG 2562
QY 2930 TCGATTTCCACACATATTTGGATGAGTTGGGTGAGACAGGACCTGGAAAGGCTTTGAGT 2989
DB 2563 GGGCACTGTGGTGGAAACGCTAGACAAGGAGATGCAAGTATTTGGTGTAGTCGGATTGCAA 2622
QY 2990 GGATTTGGAGAAATTCATCCAGATAGCAGTATTAATATGCGCGCTCTCTAAGGATA 3049
DB 2623 G--TTGGATGGCCACTTACCTGAATGACCACTTAGAGCTTTGGATCCAGGAGAACGGCGG 2680

QY 3050 GATTTACAATATCGGAGACAACGCCAAGAACACATTTGTTCTGCAATATGGACGCTGA 3109
DB 2681 CTGGAGACATCTTTTGGATCTCTACGGGAACAATACAGACCCGAGACGCGAAGAGCCA 2740
QY 3110 GACCCGAAGACACCGGGGTCTATTTTGTGAAGCTTTACTTTCGGCTTCCCTGTTTG 3169
DB 2741 GGAGGTTTCAACCGCTGTTCTGACGGGATGACTGTGGCTGTGTAGTTCTCTGCTGG 2800
QY 3170 CTTATTTGGGGCCAAAGGACCCCGGTACCGTCTCCAGTGTAAAGCCACACGACGCCAG 3229
DB 2801 CTCACTCTCAGTCGGAAGTACCAAGATTCGAGATATCCATCACACTCGGGCGCGGAA 2860
QY 3230 CGCGGACACCAACACCGGGGCCA-----CCATCGGTGCGACG 3271
DB 2861 TTCCGGATCCACTAGTAACGGCCCGCAGTGTGCGAATTCGCCCTCTCCCTCCCGCC 2920
QY 3272 CCCTGTCTCGCGCCAGAGGCGGTCTCGGCAGCGGGGGGGCGAGTCGACACGAGGG 3331
DB 2921 CCCTAAAGTTACTGGCCGAGCCGCTTGAATTAAGCCCGGTGTGTGTTGTCTATATGTG 2980
QY 3332 GGCT-----GGACTTCGCCCTTGGATC 3352
DB 2981 ATTTTCCACCATATTGCGTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCT 3040
QY 3353 CCAAACTCTGCTACTCTGCTGATGGAATCTCTTCTATCTATGTGTCTATCTACCTGCT 3412
DB 3041 CTTGACGAGCATCTCTAGGGGTCTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTTGAA 3100
QY 3413 TGTTCCTGAGAGTGAAGTTTCAGCAG-----AGCGCAGAGCCCGCC 3453
DB 3101 TGTGTTAGGAAGCAGTTCTCTGGAAGCTTCTTGAAGACAACACAGTCTGTATGCGAC 3160
QY 3454 CGGTACACGAGGGCCAGAACCAAGCTCTATAACAGCTCAATCTAGACGACGAGAGGAG 3513
DB 3161 CTTTTCAGGCGAGCGGAACCCCGCCCTGCGGACAGGTGCTCTGCGCCAAACCCACG 3220
QY 3514 TACGATGTTTGGACAGAGAGCTGGCGGACCCCTGAGATGGGGGGAAGCCG----- 3567
DB 3221 TGTATAAGATACACTGCAAGGGCGCAACACCCCGTGCACGCTTGTGAGTTGATAGT 3280
QY 3568 -----AGAGGAGAACCTCAGGAGGCTGTACAACTGACAGAAAGATAAGAT 3620
DB 3281 TGTGGAAGAGTCAATGCGTCTCTAAGCGTGTCAACAGGGGCTGAAGGATGCCCCA 3340
QY 3621 GCGCGAGGCTTACAG--TGAGATTGGATGAAGGC-----GAGCGCCGAGGGGCAAGGG 3674
DB 3341 GAAGTACCCCATGTTATGGAACTCTGATCTGGGGCTCGGTGCACATGCTTTACATGTG 3400
QY 3675 GCACGATGGCTTTTACAGGGTCTCAGTACAGC-CACCAAGGACACCTTACGACGCCCTTC 3733
DB 3401 TTTAGTCGAGTTTAAAAAGCTCTAGGCCCGCCGAAACACCGGGGACGTGTTTCTTTG 3460
QY 3734 ACATGCGAGGCTCCCGCTCGCTAACTCGAGCGGCGCGGATCCGGATTTAGTCCAATT 3793
DB 3461 AAAAACACGATGATGAGCTTGCCACAGGAATTCATGTGTAGCAAGCAGATCCTGAAGAAC 3520
QY 3794 TGTTAAAGACAGGATATCAGTGG--TCCAGGCTCTAGTTTGTGACTCAACAATATCACCAG 3851
DB 3521 ACCGCGCTGACGAGATCATGAGCTTCAAGGTGAACCTGGAGGGGTGTTGAAACACAC 3580
QY 3852 CTGAAGCCTATAGAGTAC-GAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAA 3910
DB 3581 GTGTTTACCATGGAGGCTGCGGCAAGGGCAACATCTCTGTTCCGCAACACGCTGTGTGAG 3640
QY 3911 AAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAGCTAGTCTTAAGTAAACGCCATT 3970
DB 3641 ATCCGCGTGAACAAAGGGCGCCCTCGCTTCGACATCTCTGAGCCCGCCCTTC 3700
QY 3971 TTGCAAGGCA-----TGGAATAATACATACTAGATATAG 4006
DB 3701 CAGTACGGCAACCCCACTTCAACAGTACCCCGAGGACATCAGGACTTCTTCTATCCAG 3760
QY 4007 AAGTTCAAGTTAGGAACAGAGAGACAGCAGAAATATATGGGCCAAACAGGATATCTG 4066

Db 3761 AGCTTCCCGCGGCTTCTGTACAGCGCACCCCTGCGCTACGAGCGCGCGCTGGTG 3820
Qy 4067 TGG-----TAAAGAGTTCCTGCCCCGCTCAGGCGCAAGACAGTTGGAACA 4112
Db 3821 GAGATCCGAGCGACATCAACCTGATCGAGGAGATGTTCTGTATCCGCGTGGAGTACAAG 3880
Qy 4113 GGAGAAATATGGGCCAACAGGATATCTGTGTAAGCAGTTCTGCCCCGGCTCAGGGCCA 4172
Db 3881 GCGCGCACTTCCCAACAGCGGCCCGGATGTAAGAGACCATCACCGGCCCTGAGCCC 3940
Qy 4173 AGAACAGATGTCCTCCAGATCGGGTCCCGCCCTCAGCAGTTTCTAGAGAAC----- 4223
Db 3941 AGCTTCGAGGTGGTGATCATGAACGAGCGGCTGTGTGGGCCAGTGATCTCTGTGTATC 4000
Qy 4224 -----CATCAGATGTTTCCAGGTGCGCCCAAGGACCTGAATGAC 4263
Db 4001 CGCTGAACAGCGGCAAGTTCTACAGCTGCCACATGCGCACCTGATGAAGAGCAAGGCG 4060
Qy 4264 CCGTGCCTTATTTGAACATA--ACCAATCAGTTTCGCTTCTCGCTTCTGTTCGCGGCTTC 4321
Db 4061 GTGGTAAGGACTTCCCCAGGTACCACTTCACTCCAGCACCGCTGGAGAAGACTTACGTG 4120
Qy 4322 TGCTCCCGAGCTCAATAAAGAGGCCCAACCCCTCACTCGGCGC-----GC 4369
Db 4121 GAGGACGGCGCTTCGTGGAGCAGCAGAGACGGCCATCGCCAGCTGACCGCTGGGC 4180
Qy 4370 CAGTCTCCGATAGACTCGCTCGCCGGGTACCGGTGTTCTCAATAAACCTTCTGCAGT 4429
Db 4181 AAGCCCTCGGCGAGCTGACAGATGGGTGTAACTCGAGCGCGCGGATCCGAGTTCTTTC 4240
Qy 4430 TGCATCCGACTCGTGTCTCGTGTCTTGGGAGGGTCTCT-----CTGA 4475
Db 4241 TGAGCGGAGCTTGGGGTTCGATAAATAAAGATTTTATTTAGTCTCCAGAAAAAGGG 4300
Qy 4476 GTGATTGACTACCCGTCAGCGGGTCTTTCAGTTCCTCCACCTACACAGGTCTCACTAA 4535
Db 4301 GGAATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTCCAAG 4360
Qy 4536 CATTCCTGATGTGC--CGAGGAGACTCCGTCAGCCGGTTTTTTGTTATAATAAATGCA 4593
Db 4361 GCATGAAAAATACATAAATGAGAAATAGAGAAATTCAGATCAAGGTGAGGAAACAGATGA 4420
Qy 4594 AGAACAGTCTTCCCTTCAAGCAGACTACATCTGACTCTCGGCTTTATAAAGATGTT 4653
Db 4421 ACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCTTGCCCCGGCTCAGGG 4480
Qy 4654 GAAGGGCTCTGTGGACTATCTGCCACACGACTTTTAAAGATTTTATG----- 4701
Db 4481 CCAAGAACAGATGGAACAGCTGAATATGGCCAAACAGGATATCTGTGGTAAGCAGTTCC 4540
Qy 4702 CTTCTGAGATGAGGATTTAGTCAATCTATCTCTGTCGTATTTTGTGGTCTTCTCCGTATT 4761
Db 4541 TGCCCCGGCTCAGGGCCAAAGAACAGATGTTCCCAAGATCGGTCCAGCCCTCAGCAGTTT 4600
Qy 4762 TTAATTTTC----- 4770
Db 4601 CTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAATGACCTGTGCTTCA 4660
Qy 4771 -----TAGTTTGACTCCCTTCTGAGAGCAGCGGATTCGAGAGTAGTT 4815
Db 4661 TTTGAACATAACCAATCAGTTCTGCTTCTGTTCTGTTGGCGCTTCTGCTCCCCGAGCT 4720
Qy 4816 AATACTCTGAGGCGAGGCTTCTGTGAAAAAGGTTGCTCGGCTCAGGTGAGATTTTGCCA 4875
Db 4721 CAATAAAGAGGCCCAACACCCCTCACTCGGGGGCGCCAGTCTCCGATTCAGTGTGCGCC 4780
Qy 4876 TAAAAAGGGTCTGCGCCCTGTGTACAGACAGATCGGAATCTAGAGTGCATCTCAGACT 4935
Db 4781 CCGGTACCCCGTGTATCCAATAAACCCCTCTTTCAGTTGTCATCCGACTTGTGGTCTCGCTGT 4840
Qy 4936 CCCCCGGTTCGGGGCTCTGA----- 4957

Db 4841 TCCTTTGGAGGGTCTCTCTCTGAGTGAATTGACTACCCGTGAGCGGGGTCTTTTCATTGGG 4900
Qy 4958 -----TCTCAGGCGCATCTTTGCTAGAGATCTCTACGCGCGAC--GCATCGTG 5004
Db 4901 GGCTCGTCCGGGATCGGAGAGCCCTGCCCAGGACACCGACCCACACCGGAGGTAA 4960
Qy 5005 GCGGGTACCAGAGCTCGAATTCTGTAATCATGGTCTATAG-CTGTTTCTGTGTGAAATGT 5063
Db 4961 GCTGCTGCTCGCGCGTTTCGGTGTAGCGGTGAAACCTCTGACACATGCACTGCCG 5020
Qy 5084 TATCGCTCAATTTCCACACAAATACGAGCGCGAAGCATAAAGTGTAAAGCTTGGGT 5123
Db 5021 GAGACGGTCAACAGCTTGTCTGTAAAGCGATGCGCGGAGCAGCAAGCCGTGAGCGCG 5080
Qy 5124 GCCTAATCAGT-----GAGCTAACTCA 5145
Db 5081 TCAGCGGGTGTGGCGGGTGTGCGGGCGCAGCCATGACCCAGTACAGTAGCGATAGCGGA 5140
Qy 5146 CATTAATTTGCGTTGCGCTCACTGCCCCGTCTTCAGTCGCGGAAACCTGTGTCGACAGTGC 5205
Db 5141 GTGTATACCTGGCTTAACCTATGCGCATCAGAGCAGATTGTACTGAGAGTGACACCATATGC 5200
Qy 5206 ATTAATGAATCGGCCAACGCGCGGGGAGAGCGGTTCGGTATTGGGGCGCTCTTCGGCTT 5265
Db 5201 GGTGTGAATAACCGCACAGATGCGTAAAGGAGAAAAATACCGCATCAGGCGCTCTTCGCTT 5260
Qy 5266 CCGCTCACTGACTCGCTGCGCTCGGTCTGGTCTGGCGAGCGGTATCAGTCACT 5325
Db 5261 CCGCTCACTGACTCGCTGCGCTCGGTCTGGTCTGGCGAGCGGTATCAGTCACT 5320
Qy 5326 CAAAGCGCGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGAGGAAAGAAACATGTGAG 5385
Db 5321 CAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAAACGAGGAAAGAAACATGTGAG 5380
Qy 5386 CAAAGGCGCAGAAAGGCGCAGGAAACCGTAAAAAGCGCGTGTGCGGTTTTTCCATA 5445
Db 5381 CAAAGGCGCAGAAAGGCGCAGGAAACCGTAAAAAGCGCGTGTGCGGTTTTTCCATA 5440
Qy 5446 GCGTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGAAACC 5505
Db 5441 GCGTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGAAACC 5500
Qy 5506 CGACAGGACTATAAAGATACGAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTG 5565
Db 5501 CGACAGGACTATAAAGATACGAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTG 5560
Qy 5566 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGCGC 5625
Db 5561 TTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGCGC 5620
Qy 5626 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAGCTGG 5685
Db 5621 TTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAGCTGG 5680
Qy 5686 GCTGTGTGACGAAACCCCGCTTACGCGCGACCGCTGCGCTTATCCCGTAACTATCGTC 5745
Db 5681 GCTGTGTGACGAAACCCCGCTTACGCGCGACCGCTGCGCTTATCCCGTAACTATCGTC 5740
Qy 5746 TTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGCGAGCAGCACTGTGTAAACAG 5805
Db 5741 TTGAGTCCAAACCCGTTAAGACACGACTTATCGCCACTGCGAGCAGCACTGTGTAAACAG 5800
Qy 5806 TTGAGCAGCGAGGTATGTAGCGGTGCTACAGGTTCTTGAAGTGTGGTCCCTAACTACG 5865
Db 5801 TTGAGCAGCGAGGTATGTAGCGGTGCTACAGGTTCTTGAAGTGTGGTCCCTAACTACG 5860
Qy 5866 GCTACACTAGAAGGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACTCTTCGAA 5925
Db 5861 GCTACACTAGAAGGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACTCTTCGAA 5920
Qy 5926 AAGAGTTGGTGTAGCTTTGATCCGCGAAACAAACCAACCGCTGTTAGCGGTGTTTTTTT 5985
Db 5921 AAGAGTTGGTGTAGCTTTGATCCGCGAAACAAACCAACCGCTGTTAGCGGTGTTTTTTT 5980

Db 1655 ACGAGCCATAGATAGAAATAAAGATTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAG 1714
QY 3928 ACCCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAGCCCAT 3969
Db 1715 ACCCCACCTGTAGTTTGGCAAGCTAGATCGCTTAGCTGATAGCCGAGTAAGCCCAT 1774
QY 3970 TTTGCAAGGCATGG-AAAAATACATAA-CTGAGAATAGAGAAGTTTCAGATCAAGGTTTAGGAA 4028
Db 1775 TTTGCAAGGCATGGAAAAATACCAACCAAGAATAGGAAGTTTCAGATCAAGGGGGGTA 1834
QY 4029 CAGAGAGACAGCAGAATATGGGCAAAACAGGATATCTGTGTAAGCAGATTCCTGCCCCGC 4088
Db 1835 CATGAAATAGCTTAACGTTTGGGCAAAACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGG 1894
QY 4089 TCAGGGCCAGACACGCTCGAACAGGAGATATGGGCCAAAC-----AGGATATCTG 4140
Db 1895 CCGGGGCAAGAACAGATGTCACCGCAGTTTCGGCCCCGGGCCCGGCCCAAGAACAGATG 1954
QY 4141 TGGTAAGCAGTTTCTCTGCCCGGCTCAGGGCCCAAGAACAGATGGTCCCCAGATGCGGTCCC 4200
Db 1955 GTCACCGCAGTTTCGGCCCCGGGCCCGGCCCAAGAACAGATGGTCCCCAGATGAGCCCA 2014
QY 4201 GGCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAT 4260
Db 2015 TCCCTCAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCCAAGGACCTGAAT 2074
QY 4261 GACCTGTGCTTATTTGAACCTAACCAATCAGTTGCTTCTGCTTCTGCTTCTGCTGCGGCTT 4320
Db 2075 GACCTGTGCTTATTTGAATTAACCAATCAGCTTCTGCTTCTGCTTCTGCTTCTGCTGCGGCTT 2134
QY 4321 CTGCTCCCGGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCTCCGA 4380
Db 2135 CTGCTCCCGGAGCTCTAATAAGAGCTCAACCCCTCACTCGGGCGCCAGTCTCTCCGA 2194
QY 4381 TAGACTGCTGCGCCGGGTACCGGTGTTCTCAATAAACCTCTTGCAAGTTGATCGACT 4440
Db 2195 TTGACTGAGTCGCGCGGTACCGGTGTTCTCAATAAACCTCTTGCAAGTTGATCGACT 2254
QY 4441 CGTGTCTGCTGCTTCTTGGGAGGGTCT-CTCTGAGTCAATGACTACCCGTCAGC-GGG 4498
Db 2255 CGTGTCTGCTGCTTCTTGGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCGGG 2314
QY 4499 GTCTTTTCACTTCTCCCACTACACAGCTCTCACTAAACATCTCTGATGTCGCGCAGGGAC 4558
Db 2315 GTCTTTCACTTCTCCCACTACACAGCTCTCACTAAACATCTCTGATGTCGCGCAGGGAC 2374
QY 4559 TCGTCAGCCCGGTTTTGTTATATAAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGA 4618
Db 2375 TCCGTCAAGCCCGGTTTGTGTTATATAAATAAATGCAAGAACAGTGTTCCTTCAAGCCAGA 2434
QY 4619 CTACATCCTGACTCTCGGCTTTATAAAGAATGTTGAAGGCTCTGTGGAATCTTGCCA 4678
Db 2435 CTACATCCTGACTCTCGGCTTTATAAAGAATGTTGAAGGCTCTGTGGAATCTTGCCA 2494
QY 4679 CACGACTTTTAAAGATTTTATGCTCTGAGTGAAGGATTTAGTCAATCTATCTCTGTC 4738
Db 2495 CACGAC-TTTTAAAGATTTTATGCTCTGAGTGAAGGATTTAGTCAATCTATCTCTGTC 2553
QY 4739 TATTTGCTGCTTCTCCGTAATTTTAAATTTCTAGTTTCCACTCCCTTCTGAGACAG 4798
Db 2554 TATTTGCTGCTTCTCCGTAATTTTAAATTTCTAGTTTTCGACTCCCTTCTGAGACAG 2613
QY 4799 GCGATTGCAAGTAGTAACTCTGAGGCGAGGCTTCTGTGAAGGTTGCTGGGCTC 4858
Db 2614 GCGATTGCAAGTAGTAACTCTGAGGCGAGGCTTCTGTGAAGGTTGCTGGGCTC 2673
QY 4859 AGTGTGAGATTTTGCATAAAGAGGCTCTGCGCCCTGTGTACAGACAGATCGGAATCTA 4918
Db 2674 AGTGTGAGATTTTGCATAAAGAGGCTCTGCGCCCTGTGTACAGACAGATCGGAATCTA 2733
QY 4919 GAGTGATATCAGAGTCCCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTAG 4978
Db 2734 GAGTGATATCAGAGTCCCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGCTAG 2793

QY 4979 AGATCCTCTACGCCGACGATCGTGGCGGGTACCGAGCTCGAATTCGTATCATGTGTC 5038
Db 2794 AGATCCTCTACGCCGACGATCGTGGCGGGTACCGAGCTCGAATTCGTATCATGTGTC 2853
QY 5039 ATAGCTGTTTC-----CTGTGTGAATTTGTTATCCGCTCAAAATTCACACA 5085
Db 2854 GGGCTTATATCGCCGACATCACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATGA 2913
QY 5086 ACATACGAGCGGAGCATATAAGTGTAAAGCTTGGGTGGCTTAATAGT----- 5134
Db 2914 GCGCTTGTTCGGCTGGGTATGTTGGCAGGCCCTGTCGGGGGACTGTTTGGGCGCA 2973
QY 5135 -----GAGCTAACTCACATTAATTCGTTGCGC 5162
Db 2974 TCTCCTTGATGACACCATTCCTTGGCGCGGGTCTCAACGGCTCAACCTACTACTGG 3033
QY 5163 TCATTCGCCGCTTTCAGTTCGGGAAACCTGTGTGCCAGCTGCAATTAATGAATCGGCCAA 5222
Db 3034 GCTGCTTCTTAATGACGAGTCCGATAAGGAGAGCGTCTGCAATTAATGAATCGGCCAA 3093
QY 5223 CGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTTCGGCTTCTCGCTCACTACTCG 5282
Db 3094 CGCGGGGAGAGCGGTTTGGGTATTTGGGCGCTTTCGGCTTCTCGCTCACTACTCG 3153
QY 5283 CTGCGCTCGGTCGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGG 5342
Db 3154 CTGCGCTCGTTCGTCGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAATACGG 3213
QY 5343 TTATCCACAGATCAGGGGATAACACGAGGAAGAAACATGTGAGCAAAAGGCCAGCAAAAG 5402
Db 3214 TTATCCACAGATCAGGGGATAACACGAGGAAGAAACATGTGAGCAAAAGGCCAGCAAAAG 3273
QY 5403 GCAGGAACCGTAAAGGCGGTTGCTGGGCTTTTCCATAGCTCGGCCCTGAC 5462
Db 3274 GCAGGAACCGTAAAGGCGGTTGCTGGGCTTTTCCATAGCTCGGCCCTGAC 3333
QY 5463 GAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGA 5522
Db 3334 GAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGA 3393
QY 5523 TACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGCTT 5582
Db 3394 TACAGGGGTTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCGCTT 3453
QY 5583 ACCGATACGTCGCTTCTCCCTCGGGAAGCGTGGGCTTTCATAGCTCAGCG 5642
Db 3454 ACCGATACGTCGCTTCTCCCTCGGGAAGCGTGGGCTTTCATAGCTCAGCG 3513
QY 5643 TGTAGTATCTCAGTTCGCTGCTGCTCGCTCAAGCTGGGCTGTGTCAGCAACCC 5702
Db 3514 TGTAGTATCTCAGTTCGCTGCTGCTCGCTCAAGCTGGGCTGTGTCAGCAACCC 3573
QY 5703 CCGCTTCAGCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTA 5762
Db 3574 CCGCTTCAGCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCGGTA 3633
QY 5763 AGACAGACTTATCCGCTGCGCAGCAGCTGCTGTAACAGGATTTAGCAGCGAGGAT 5822
Db 3634 AGACAGACTTATCCGCTGCGCAGCAGCTGCTGTAACAGGATTTAGCAGCGAGGAT 3693
QY 5823 GTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTCGGCTACACTAGAAAGACA 5882
Db 3694 GTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTCGGCTACACTAGAAAGACA 3753
QY 5883 GTATTTGTTATCTGCGCTCTGCTGAAGCAGGTTACTCTCGGAAAAAGAGTTGGTAGCTCT 5942
Db 3754 GTATTTGTTATCTGCGCTCTGCTGAAGCAGGTTACTCTCGGAAAAAGAGTTGGTAGCTCT 3813
QY 5943 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTTTTGTTTGCAGCAGAGATT 6002
Db 3814 TGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTTTTGTTTGCAGCAGAGATT 3873

Qy	6003	ACGGCGACAAAAAGGATCTCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCT	6063
Db	3874	ACGGCGACAAAAAGGATCTCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCT	3933
Qy	6063	CAGTGGACAGAAAACCTACAGTTAAAGGATTTTGGTCATGAGATTAACAAGAGATCTTC	6122
Db	3934	CAGTGGACAGAAAACCTACAGTTAAAGGATTTTGGTCATGAGATTAACAAGAGATCTTC	3993
Qy	6123	ACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAA	6182
Db	3994	ACCTAGATCTCTTTTAAATTTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAA	4053
Qy	6183	ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTCCTA	6242
Db	4054	ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTCCTA	4113
Qy	6243	TTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCCGTGTAGATAACTACGATACGGGAGGC	6302
Db	4114	TTTCGTTTCATCCATAGTTTGCCTGACTCCCGTCCGTGTAGATAACTACGATACGGGAGGC	4173
Qy	6303	TTACCATCTGGCCCCAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT	6362
Db	4174	TTACCATCTGGCCCCAGTGTCTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGAT	4233
Qy	6363	TTATCAGCAATAAACAGCAGCCAGCCGGAAGGGCCGAGCGAGAAGTGTCTGCAACTTTA	6422
Db	4234	TTATCAGCAATAAACAGCAGCCAGCCGGAAGGGCCGAGCGAGAAGTGTCTGCAACTTTA	4293
Qy	6423	TCCGCCCTCCATCCAGTCTATTAATGTGTTCGGGAAGCTAGATTAAGTATTGCCAGTT	6482
Db	4294	TCCGCCCTCCATCCAGTCTATTAATGTGTTCGGGAAGCTAGATTAAGTATTGCCAGTT	4353
Qy	6483	AATAGTTTTCGCAACAGTTTGTGCCATTCGTACAGGC--TCGTGGTGTCAACGCTCGTCGTT	6541
Db	4354	AATAGTTTTCGCAACAGTTTGTGCCATTCGTACAGGC--TCGTGGTGTCAACGCTCGTCGTT	4413
Qy	6542	GGTATGCTTCANTCAGTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATG	6601
Db	4414	GGTATGCTTCANTCAGTCCGGTTCCCAACGATCAAGCGAGTTACATGATCCCCCATG	4473
Qy	6602	TTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTGTCAGAACTAAGTTGGCC	6661
Db	4474	TTGTGCAAAAAGCGGTAGCTCCTTCGGTCTCCGATCGTTGTCAGAACTAAGTTGGCC	4533
Qy	6662	GCAGTGTATCACTCATGGTTATGGCAGCACTGCATAAATCTCTTACTGTCATGCCATCC	6721
Db	4534	GCAGTGTATCACTCATGGTTATGGCAGCACTGCATAAATCTCTTACTGTCATGCCATCC	4593
Qy	6722	GTAAGATGCTTTTCTGTCAGTGTGAGTACTCAACCAAGTCAITCTGAGAAATAGTGTATG	6781
Db	4594	GTAAGATGCTTTTCTGTCAGTGTGAGTACTCAACCAAGTCAITCTGAGAAATAGTGTATG	4653
Qy	6782	CGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATTAATACCGCGCCACATAGCAGA	6841
Db	4654	CGGCGACCGAGTTGCTCTTGGCCGGCGTCAATACGGGATTAATACCGCGCCACATAGCAGA	4713
Qy	6842	ACTTTAAAGTGCTCATCTATTGGAAAAAGTTCTTTCGGGGCGAAAACTCTCAAGAGTCTTA	6901
Db	4714	ACTTTAAAGTGCTCATCTATTGGAAAAAGTTCTTTCGGGGCGAAAACTCTCAAGAGTCTTA	4773
Qy	6902	CCGCTGTTGAGATCCAGTTGATGTAAACCACTGTCACCCAACTGATCTTCAGCATCT	6961
Db	4774	CCGCTGTTGAGATCCAGTTGATGTAAACCACTGTCACCCAACTGATCTTCAGCATCT	4833
Qy	6962	TTTACTTTTCCACCGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAG	7021
Db	4834	TTTACTTTTCCACCGGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAAAAAG	4893
Qy	7022	GGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATATTGA	7081
Db	4894	GGAATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATATTGA	4953
Qy	7082	AGCATTTTATCAGGGTTATTGTCTCATGAGCGGATACATAATTGAATGTAATTTAGAAAAAT	7141

Db	4954	AGCATTTATCAGGGTTATTGCTCTCATGAGCGGATACATATTGTAATGTATTTAGAAAAAT	5013
Qy	7142	AAACAAATAGGGTTCCGGCGACATTTCCCGGAAAGTGCACCTGACGCTCTAAGAAACC	7201
Db	5014	AAACAAATAGGGTTCCGGCGACATTTCCCGGAAAGTGCACCTGACGCTCTAAGAAACC	5073
Qy	7202	ATTATTATCATGACATTAAACCTATAAAAAATAGCGGTATCACAGAGGCCCTTTTCGTCGCG	7261
Db	5074	ATTATTATCATGACATTAAACCTATAAAAAATAGCGGTATCACAGAGGCCCTTTTCGTCGCG	5133
Qy	7262	CGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAAGCT	7321
Db	5134	CGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAAGCT	5193
Qy	7322	TGTCGTGAACGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGGC	7381
Db	5194	TGTCGTGAACGGATGCCGGGAGCAGACAAGCCCGTCAGGCGCGTCAGCGGGTGTGGC	5253
Qy	7382	GGGTGTCCGGGGCTGGCTTAACTATGCGGCATCAGACAGANTTCTATGAGAGTGCACCAT	7441
Db	5254	GGGTGTCCGGGGCTGGCTTAACTATGCGGCATCAGACAGANTTCTATGAGAGTGCACCAT	5313
Qy	7442	ATCGGTGTGAAATACCGCACAGATCGGTGAAGAGAAAAATACCGCATCAGG	7492
Db	5314	ATCGGTGTGAAATACCGCACAGATCGGTGAAGAGAAAAATACCGCATCAGG	5364

RESULT 12

AAI67595
ID AAI67595 standard; DNA; 6444 BP.

XX
AC
AAI67595:

XX DT 27-FEB-2002 (first entry)

XX DE Nucleotide sequence of a cloning vector pLXRN.

XX	Hexokinase; cell proliferation; glycolytic tumour; cancer; mutant;
KW	type II hexokinase; cytostatic; glycolysis inhibitor; gene therapy; da.
XX	Synthetic.
OS	

xx PT Inhibiting proliferation of highly glycolytic tumors, e.g. in gastric
PT cancer, hepatoma, colorectal cancer or lung cancer, by contacting cells
PT with antisense molecules that hybridize with a nucleic acid encoding
PT hexokinase.

XX PS Claim 21; Fig 7A-C; 59pp; English.

xx The invention provides a method for inhibiting proliferation of tumour
 CC cells characterized by having a highly glycolytic phenotype. The method
 CC involves contacting the cells with an antisense polynucleotide or
 CC oligonucleotide that hybridizes with a mRNA encoding a hexokinase under
 CC conditions that allow hybridization of the antisense polynucleotide with
 CC the mRNA, thus inhibiting the proliferation of tumour cells. The method
 CC is useful for inhibiting proliferation of highly glycolytic tumours or
 CC for modulating the expression of a hexokinase in highly glycolytic
 CC tumours. In particular, the cellular proliferative disorder comprises low

CC grade astrocytoma, anaplastic astrocytoma, glioblastoma, medulloblastoma,
CC gastric cancer, hepatoma, colorectal cancer, colorectal adenoma, acute
CC myelogenous leukemia, lung cancer, renal cancer, leukemia, breast cancer,
CC prostate cancer, endometrial cancer, bone cancer, squamous cell cancer,
CC and neuroblastoma. The present sequence represents the nucleotide
C sequence of a cloning vector pLXRN

Sequence 6444 BP; 1476 A; 1758 C; 1668 G; 1542 T; 0 U; 0 Other;

Query Match 40.5%; Score 3103.4; DB 6; Length 6444;

Best Local Similarity 71.0%; Pred. No. 0;

Sequence	Matches	Mismatches	Indels	Gaps
Conservative	4779	0	1466	486
Similarity	4779	0	1466	486
Identity	4779	0	1466	486

QY 549 TTTTGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGG 608

db 174 TTTTGAAGACCCCA CCGTAGG-- TGGCAAGCTAGCTTAAGTAA CGCCACCTTTTGCAAGG 231

609 CATGG-AAAATACATAA CTGAGGAATAGAGAGAAGTTCAATCAAGGTTAGGACACAGAGAGAC 667

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

.....

Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14, Q15, Q16, Q17, Q18, Q19, Q20, Q21, Q22, Q23, Q24, Q25, Q26, Q27, Q28, Q29, Q30, Q31, Q32, Q33, Q34, Q35, Q36, Q37, Q38, Q39, Q40, Q41, Q42, Q43, Q44, Q45, Q46, Q47, Q48, Q49, Q50, Q51, Q52, Q53, Q54, Q55, Q56, Q57, Q58, Q59, Q60, Q61, Q62, Q63, Q64, Q65, Q66, Q67, Q68, Q69, Q70, Q71, Q72, Q73, Q74, Q75, Q76, Q77, Q78, Q79, Q80, Q81, Q82, Q83, Q84, Q85, Q86, Q87, Q88, Q89, Q90, Q91, Q92, Q93, Q94, Q95, Q96, Q97, Q98, Q99, Q100, Q101, Q102, Q103, Q104, Q105, Q106, Q107, Q108, Q109, Q110, Q111, Q112, Q113, Q114, Q115, Q116, Q117, Q118, Q119, Q120, Q121, Q122, Q123, Q124, Q125, Q126, Q127, Q128, Q129, Q130, Q131, Q132, Q133, Q134, Q135, Q136, Q137, Q138, Q139, Q140, Q141, Q142, Q143, Q144, Q145, Q146, Q147, Q148, Q149, Q150, Q151, Q152, Q153, Q154, Q155, Q156, Q157, Q158, Q159, Q160, Q161, Q162, Q163, Q164, Q165, Q166, Q167, Q168, Q169, Q170, Q171, Q172, Q173, Q174, Q175, Q176, Q177, Q178, Q179, Q180, Q181, Q182, Q183, Q184, Q185, Q186, Q187, Q188, Q189, Q190, Q191, Q192, Q193, Q194, Q195, Q196, Q197, Q198, Q199, Q200, Q201, Q202, Q203, Q204, Q205, Q206, Q207, Q208, Q209, Q210, Q211, Q212, Q213, Q214, Q215, Q216, Q217, Q218, Q219, Q220, Q221, Q222, Q223, Q224, Q225, Q226, Q227, Q228, Q229, Q230, Q231, Q232, Q233, Q234, Q235, Q236, Q237, Q238, Q239, Q240, Q241, Q242, Q243, Q244, Q245, Q246, Q247, Q248, Q249, Q250, Q251, Q252, Q253, Q254, Q255, Q256, Q257, Q258, Q259, Q260, Q261, Q262, Q263, Q264, Q265, Q266, Q267, Q268, Q269, Q270, Q271, Q272, Q273, Q274, Q275, Q276, Q277, Q278, Q279, Q280, Q281, Q282, Q283, Q284, Q285, Q286, Q287, Q288, Q289, Q290, Q291, Q292, Q293, Q294, Q295, Q296, Q297, Q298, Q299, Q300, Q301, Q302, Q303, Q304, Q305, Q306, Q307, Q308, Q309, Q310, Q311, Q312, Q313, Q314, Q315, Q316, Q317, Q318, Q319, Q320, Q321, Q322, Q323, Q324, Q325, Q326, Q327, Q328, Q329, Q330, Q331, Q332, Q333, Q334, Q335, Q336, Q337, Q338, Q339, Q340, Q341, Q342, Q343, Q344, Q345, Q346, Q347, Q348, Q349, Q350, Q351, Q352, Q353, Q354, Q355, Q356, Q357, Q358, Q359, Q360, Q361, Q362, Q363, Q364, Q365, Q366, Q367, Q368, Q369, Q370, Q371, Q372, Q373, Q374, Q375, Q376, Q377, Q378, Q379, Q380, Q381, Q382, Q383, Q384, Q385, Q386, Q387, Q388, Q389, Q390, Q391, Q392, Q393, Q394, Q395, Q396, Q397, Q398, Q399, Q400, Q401, Q402, Q403, Q404, Q405, Q406, Q407, Q408, Q409, Q410, Q411, Q412, Q413, Q414, Q415, Q416, Q417, Q418, Q419, Q420, Q421, Q422, Q423, Q424, Q425, Q426, Q427, Q428, Q429, Q430, Q431, Q432, Q433, Q434, Q435, Q436, Q437, Q438, Q439, Q440, Q441, Q442, Q443, Q444, Q445, Q446, Q447, Q448, Q449, Q450, Q451, Q452, Q453, Q454, Q455, Q456, Q457, Q458, Q459, Q460, Q461, Q462, Q463, Q464, Q465, Q466, Q467, Q468, Q469, Q470, Q471, Q472, Q473, Q474, Q475, Q476, Q477, Q478, Q479, Q480, Q481, Q482, Q483, Q484, Q485, Q486, Q487, Q488, Q489, Q490, Q491, Q492, Q493, Q494, Q495, Q496, Q497, Q498, Q499, Q500, Q501, Q502, Q503, Q504, Q505, Q506, Q507, Q508, Q509, Q510, Q511, Q512, Q513, Q514, Q515, Q516, Q517, Q518, Q519, Q520, Q521, Q522, Q523, Q524, Q525, Q526, Q527, Q528, Q529, Q530, Q531, Q532, Q533, Q534, Q535, Q536, Q537, Q538, Q539, Q540, Q541, Q542, Q543, Q544, Q545, Q546, Q547, Q548, Q549, Q550, Q551, Q552, Q553, Q554, Q555, Q556, Q557, Q558, Q559, Q560, Q561, Q562, Q563, Q564, Q565, Q566, Q567, Q568, Q569, Q570, Q571, Q572, Q573, Q574, Q575, Q576, Q577, Q578, Q579, Q580, Q581, Q582, Q583, Q584, Q585, Q586, Q587, Q588, Q589, Q590, Q591, Q592, Q593, Q594, Q595, Q596, Q597, Q598, Q599, Q600, Q601, Q602, Q603, Q604, Q605, Q606, Q607, Q608, Q609, Q610, Q611, Q612, Q613, Q614, Q615, Q616, Q617, Q618, Q619, Q620, Q621, Q622, Q623, Q624, Q625, Q626, Q627, Q628, Q629, Q630, Q631, Q632, Q633, Q634, Q635, Q636, Q637, Q638, Q639, Q640, Q641, Q642, Q643, Q644, Q645, Q646, Q647, Q648, Q649, Q650, Q651, Q652, Q653, Q654, Q655, Q656, Q657, Q658, Q659, Q660, Q661, Q662, Q663, Q664, Q665, Q666, Q667, Q668, Q669, Q670, Q671, Q672, Q673, Q674, Q675, Q676, Q677, Q678, Q679, Q680, Q681, Q682, Q683, Q684, Q685, Q686, Q687, Q688, Q689, Q690, Q691, Q692, Q693, Q694, Q695, Q696, Q697, Q698, Q699, Q700, Q701, Q702, Q703, Q704, Q705, Q706, Q707, Q708, Q709, Q710, Q711, Q712, Q713, Q714, Q715, Q716, Q717, Q718, Q719, Q720, Q721, Q722, Q723, Q724, Q725, Q726, Q727, Q728, Q729, Q730, Q731, Q732, Q733, Q734, Q735, Q736, Q737, Q738, Q739, Q740, Q741, Q742, Q743, Q744, Q745, Q746, Q747, Q748, Q749, Q750, Q751, Q752, Q753, Q754, Q755, Q756, Q757, Q758, Q759, Q760, Q761, Q762, Q763, Q764, Q765, Q766, Q767, Q768, Q769, Q770, Q771, Q772, Q773, Q774, Q775, Q776, Q777, Q778, Q779, Q780, Q781, Q782, Q783, Q784, Q785, Q786, Q787, Q788, Q789, Q790, Q791, Q792, Q793, Q794, Q795, Q796, Q797, Q798, Q799, Q800, Q801, Q802, Q803, Q804, Q805, Q806, Q807, Q808, Q809, Q810, Q811, Q812, Q813, Q814, Q815, Q816, Q817, Q818, Q819, Q820, Q821, Q822, Q823, Q824, Q825, Q826, Q827, Q828, Q829, Q830, Q831, Q832, Q833, Q834, Q835, Q836, Q837, Q838, Q839, Q840, Q841, Q842, Q843, Q844,

DD
 348 AAGAAACAGATGAGACACAGCTGAGTGTATGGCCATATACAGGATATCTGTGTATAGCAGTTCCT 407

QY 786 GCCCGGCTCAGGGCCCAAGAACAGATGGTCCCCCAGATGCGGTCCCCGCCCTCAGCAGTTTC 845
QY

Db 408 GCCCGGCTCGGGGCCAAGAACAGATGGTGTCCCAAGCTCAGCCCTCAGCAGTTTC 467

Qy 846 TAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTG-AAATGACCCCTGTGCCTTA 904

Db 468 TAGTGAATCATCAGATGTTTCCAGGTGCCCCAAGGACCTGAAAATGACCCCTGTACCTTA 527

QY 905 TTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTTCGGCGGCTTCTGCTCCCCGAGCT 964

Db 528 TTTGAAC^TAACCAATCAGTT^CCGCTTCTCGCTTCTGTT^CCGCGGCTTCCGAGCT 587

QY 965 CAATAAAGAGCCCAACAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGGGTGGC 102

db 588 CAATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTTCCGATAGACTGGCTCGCC 647

1025 CCGGTACCCGTAATTCCCAATAAAGCCTCTTGCTGTTGCA TCCGAATCGTGGA CTGGCTG 108

b6
648 CCGCTACCCGATATTCCCATAAAGCCTTCTGTTTCATCCCCATCCTGCCTGCCCTC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

DD 788 GGCTCGTCGGGATITGGAGACCCCTGTCCAGGGGACCACCGA-CCCAACCCGGAGGGA 828

1205 AGCTGGCCAGCAACTTATCTGTGCTGTCCGATGTCTAGTGTCTATGACCTGATTTTATG 1264

Db 827 AGCTGGCCAGCAACTTATCTGTGTCTGTGCCGATTGTCTAGTGTCTATGTTTGATGTTATG 886

Qy 1265 CGCCTGCGTACCTAGTAACTAGCTCTGTATCTGGCGGACCCGTTGGTGGAACT 1324

Db 887 CGCCTGCGTCTGTACTAGTAACTAGCTCTGTATCTGGCGGACCCGTTGGTGAAC 946

QY 1325 GACGAGTTCGGAACACCGCGCAACCTGGGAGACGTCCAGGACTTCGGGGGCCGT 1384

Db
947 GACGAGTTCTGAACACCCGGCCGCAACCTGGGAGACGTCCAGGGACTTTGGGGGCCGT 1006

Db 2122 CAATGTAGGGCCCGGAAACCTGGCCCTGTCTTCTTTGAGGAGCATTCCTAGGGGCTTTC 2181
Qy 2580 TGTAGCTTGGTACACGAGAGAGCAGGTAAGGCTCCAAAGCTGTGATCTACTTGGACATC 2639
Db 2182 -----CCCTCTCGCCCAAGAAATGCAAGGCTCTGTGAATGCTCGTGAAGGAGCAGTTC 2234
Qy 2640 CACCCGGCACACTGTGTGCGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCAC 2699
Db 2235 CTCTGGAAGCTTCTTTGAAGACAAACAAAGCTCTGTAGCAGCCCTTTGCAAGGAGCGGAACC 2294
Qy 2700 CTTCACCATCAGCAGCCTCCAGCCAGAGAGACATCCGACCTACTACTGCCAGCAATATAG 2759
Db 2295 CCCCACTGGCGACAGGTGCCTCTCGCGCCAAAGCCAGCT-----GTATAGATACA 2347
Qy 2760 CCTTATCGGTGCTTGGCCAAAGGACCAAGGTGGAATCAAAAGAGGTGCTCAGGATC 2819
Db 2348 CCTGCAAAAGGGGACAAACCCAGGTGCCAGCTGTGTGAGTTGAGTTGTTGGAAGAGATC 2407
Qy 2820 GGGTGGATCCGGCTCTGTGGCTCAGGATCGAGGTCCAATCTGTGGAGAGCGGTGGAGG 2879
Db 2408 AAATGGCTCTCCTCAAGCGTAGTCAACAAGGGGCTGAAGGATGCCAGAGGTATCCCOAT 2467
Qy 2880 TGTGTGCAACTGGCCGCTCCCTCGCTGTCTGCTCCGATCTGCTGCTTTCGATTTCCAC 2939
Db 2468 TGTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGT-----TTTAG 2519
Qy 2940 CACATATTTGGATGAGTTGGGTGAGACAGGCACCTGGAAGAGTCTTTGAGTGGATTTGGAGA 2999
Db 2520 TCGAGTTAAAGAGCTCTAGGCCCCCGAACCACGGGACGTGGTTTCTTT-----GA 2575
Qy 3000 AATTTCATCCAGATAGCAGTACGATTAATCTATGCGCCGTCTCTAAAGGATGATTTACAAT 3059
Db 2576 AAAACACATGATAGCTTTGCCAACAAACCCCGGATAATTCCTGCAAGCAATATGGGATCG 2635
Qy 3060 ATCGCGACACGCCAAGAGACATTTCTCTGCAATGACGACCTGACACCGA--A 3117
Db 2636 GCCATTGAACAAGATGGATTGACGACGAGTTCTCCGGCCGCTTGGGTGGAGGCGGTATTC 2695
Qy 3118 GACACCGGGTCTATTTTGTGCAAGCCTTTACTTTGGCTTCCCTGGTTTGTCTTATTGG 3177
Db 2696 GGCTATGACTGGGCACACAGACATCGGTCTCTGATGCGCGCTGTTCGGCTGTCA 2755
Qy 3178 GGCAAGGAGACCCGGTCAACGTCTCCAGTGTCAAGCCCAACAGACGCCAGCGCCGGA 3237
Db 2756 GCGAGGGGCGCCGGTCTTTTGTCAAGACCGACCTGTCCGGTGCC-----TGAATGAA 2812
Qy 3238 CCACCAACACCGGCGCCACCATCGCGTCGAGCCCTGTCCCTGCGCCAGAGCGGCT 3297
Db 2813 CTGACGAGGAGGACGCGCGGTATCGTGGCTGGCCACGACGCGGCGTTCTTGGCGAGCT 2872
Qy 3298 CGGCGAGCGGGGGGCGAGTGACACAGAGGGGCTGGACTTCGCCCTGGATCCCAAA 3357
Db 2873 GTGCTGAGGTTGTCACTGAAGGGAAGGAGTGGCTGTATTGGGCGAAGTGGCGGG 2932
Qy 3358 CTCTGCTACTGCTGGATGGAATCCTCTTCACTATATGGTGTCACTTCACTGCTTGTTC 3417
Db 2933 CAGGATCTCTGTCACTCACTTGTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCA 2992
Qy 3418 CTGAGATGGAAGTTACGACGAGGCGCAGAGCCCGCGGTACCGAGCGGCGCAACACAG 3477
Db 2993 ATGCGGCGGTGCAATCGCTTTGATCCGGCTTACTTGCCTATTCGACCAACCAAGCGAAACAT 3052
Qy 3478 CTCTATACGAGTCAATCTTAGGACGAAGAGGAGTACGATGTTTGGCAAGAGAGCT 3537
Db 3053 CGC-----ATCGACGAGACGTACTCGGATGGAAGCGGTCTTGTGATCAGGAT 3103
Qy 3538 GCGCGGACCTGTAGATGGGGGGAAGCCGAGAGGAAGAACCTTCAGGAAGGCTGTAC 3597
Db 3104 GATCTGGACGAGAGCATCAGGG--GCTCGCGCCAGCCGACTGTTGCCAGGCTCAAGG 3161
Qy 3598 AATGACGTGCAGAAAGATAAGATGCGGAGCGCTTACAGTGAGATTGGGATGAAAGGGAG 3657

Db 3162 CGCGCATGCCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCTTGTTCGCAATA 3221
Qy 3658 CGCCGGAGGGCAAGG--GGCACGATGGCCTTTTACCAGGGTCTCAGTACAGCACCAAGG 3715
Db 3222 TCATGGTGAAATATGGCGCTTTTCTGGATTATCATGACTGTGGCGGCTGGGTGTGGCG 3281
Qy 3716 ACACCTACGACGCCCTTACATGACGAGCCCTGCCCCCTCGCTAA-----CTCGACGGGCC 3771
Db 3282 ACCGCTATCAGGACATAGCGTTGGTACCCTGTGATATTGCTGAAGAGCTTGGCGCGAAT 3341
Qy 3772 CGCGATCCGATTAGTCCAATTTGTTAAAGACAGGATATCAGTGTCTCAGGCTCTAGTTT 3831
Db 3342 GGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCGCTCCCGATTCGCGACGCTCGCT 3401
Qy 3832 TGACTCAACAATATCAACAGCTGAAGCTATAGAGTACGAGCCATAGATATAAAATAAGA 3891
Db 3402 TCTATCGCTTCTTGAACGAGTCTTCTGAGCGGACTCTGGGGTTCGATATAAAATAAGA 3461
Qy 3892 TTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGC 3951
Db 3462 TTTTATTAGTCTCCAGAAAAGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGC 3521
Qy 3952 TAGCTTAAGTAAAGCCATTTTCAAGGCATCG--AAAATACATACTAGAGATAGAGAGT 4010
Db 3522 TAGCTTAAGTAAAGCCATTTTGAAGGCATGGAAGCAATACATNACTAGAGATAGAGAGT 3581
Qy 4011 TCAGATCAAGGTTAGGAACAGA--GAGACAGAGATATGGCCCAACAGGATATCTGTGG 4069
Db 3582 TCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGG 3641
Qy 4070 TAACGAGTCTCTGCCCC--GCTCAGGGCCAGAAACAGTTGGAAACAGGAGAAATATGGGCCAA 4128
Db 3642 TAACGAGTCTCTGCCCCGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCAA 3701
Qy 4129 ACAGGATATCTGTGTAGCAGTCTCTGCCCGCTCAGGGCCAGAACAGATGCTTCCCC 4188
Db 3702 ACAGGATATCTGTGTAGCAGTCTCTGCCCGCTCAGGGCCAGAACAGATGCTTCCCC 3761
Qy 4189 AGATCGGCTCCCGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCCCA 4248
Db 3762 AGATCGGCTCCAGCCCTCAGCAGTCTTCTAGAGAACCATCAGATGTTTCCAGGGTCCCCCA 3821
Qy 4249 AGGACCTGAAATGACCTGTGCTTATTGTAATCAACATCAGTTGCTTCTCGCTTCT 4308
Db 3822 AGGACCTGAAATGACCTGTGCTTATTGTAATCAACATCAGTTGCTTCTCGCTTCT 3881
Qy 4309 GTTCGCGCTTCTGCTCCCGGCTCAATAAAGAGCCACAAACCCCTCACTCGGCGG 4368
Db 3882 GTTCGCGCTTCTGCTCCCGGCTCAATAAAGAGCCACAAACCCCTCACTCGGCGG 3941
Qy 4369 CCAGTCTCCGATAGACTGCTGCTCCCGGCTACCCGTTCTCAATAAACCCCTCTTGCAG 4428
Db 3942 CCAGTCTCCGATAGACTGCTGCTCCCGGCTACCCGTTCTCAATAAACCCCTCTTGCAG 4001
Qy 4429 TTGATCCGACTCGTGTCTGCTGCTTCTTGGGAGGGTCT--CTCTGAGTGAATGACTAC 4487
Db 4002 TTGATCCGACTCGTGTCTGCTGCTTCTTGGGAGGGTCTCTCTCTGAGTGAATGACTAC 4061
Qy 4488 CCCTCAGGGGCTTCTCAGTCTTCTCCCACTACAGGCTCTCACTAATCTCTGATGT 4547
Db 4062 CCCTCAGGGGCTTCTTCTCATTGCGGGCTCGTCCGGAGAGCCCTGCGCCAGG 4121
Qy 4548 GCCCAGGAGTCTCGTCCAGCCCGGTTTGTGTTTATATAAATGCAAGAACAGATGTTCCC 4607
Db 4122 ACACACCAACCAACCGGAGGTAAAGT----- 4151
Qy 4608 TTCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGAG 4667
Db 4152 -----GCTGCTCGCGGTTTCGGTGTATGACGCTGAAACCTCTG--AC 4193
Qy 4668 ACTATCTGCCACACGACTTTTAAAGATTTTATGCCCTCTGATGAGGATTTAGTCAAT 4727
Db 4194 ACATGACGCTCCCGAGACGCTCAAGCTTGTCTGTAAGCGGATGCCGGAGCAGACAAG 4253

QY	4728	CTATCCTCGTCTATTGTTGCTGGCTTCTCGGTATTTTAAATTTCTAGTTTGCACTCCCTTC	4787	5808	ACGACGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTCGCG	5867
Db	4254	CCCGTCAGGCGCGTTCAGCGGTTTGGCG-----	4283	4990	ACGACGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTCGCG	5049
QY	4788	CTGAGAGCAGCGGATTGAGAGTAGTTAATCTCTGAGGCGAGGTTCTGTGAAAGGT	4847	5868	TACACTAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGGTACCTTCGGAAA	5927
Db	4284	-----	4283	5050	TACACTAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCAGGTACCTTCGGAAA	5109
QY	4848	TGCTGGGCTCAGTGTGAGATTGTGCATAAAAGGGTCTTGCCCTCTGTACAGACAG	4907	5928	AGAGTTGGTAGTCTTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTTTTTTTGT	5987
Db	4284	-----	4283	5110	AGAGTTGGTAGTCTTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTTTTTTTGT	5169
QY	4908	ATCGGAATCTAGAGTGCATACTCAGAGTCCCGCGGTTCGGGGTCTGATCTCAGGGCA	4967	5988	TGCAAGCAGCAGATTACGCGCAGAAAAAAGSATCTCAAGAGATCTCTTTGATCTTTTCT	6047
Db	4284	-----GGTGTGGGCGCAGCCATGA-----	4304	5170	TGCAAGCAGCAGATTACGCGCAGAAAAAAGSATCTCAAGAGATCTCTTTGATCTTTTCT	5229
QY	4968	TCTTTGCCCTAGAGATCCTCTACGCGGACGCATCGTGGCGGGTACCGAGCTCGAATCG	5027	6048	ACGGGGTCTGAGCTCAGTGGAAACGAAATCTACGTTAAGGATTTTGGTCAATGAGATTA	6107
Db	4305	-----	4304	5230	ACGGGGTCTGAGCTCAGTGGAAACGAAATCTACGTTAAGGATTTTGGTCAATGAGATTA	5289
QY	5028	TAATCATGGCTAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACATTTCCACACAC	5087	6108	TCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAA	6167
Db	4305	-----CCAGTCAAG	4314	5290	TCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATCAATCTAA	5349
QY	5088	ATACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCCTAATGAGTGAGCTAACTCACA	5147	6168	AGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATC	6227
Db	4315	TAGCGATAGCGGNGTATATCGGCTTAATCTATGCGGATCAGACGAGATTGTTACTGAGA	4374	5350	AGTATATATGAGTAAACTTTGGTCTGACAGTTTACCAATGCTTAATCAGTGAGGCACCTATC	5409
QY	5148	TTAATTTGCGTTCGCTCACTGCCCGCTTTCCAGTGGGAAACCTGTGCGCAGCTGCAT	5207	6228	TCAGGATCTGTCTAATTTGTTGTTTCCATAGTTGCTGACTCCCGTCTGTGTAGATACT	6287
Db	4375	GT-----GCACCATATGCGG	4389	5410	TCAGGATCTGTCTAATTTGTTGTTTCCATAGTTGCTGACTCCCGTCTGTGTAGATACT	5469
QY	5208	TAATGAATCGGCAACGCGCGGGAGAGGGTTTGGCTATTGGGGCTCTTCCGCTTCC	5267	6288	ACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGAGACCCACGC	6347
Db	4390	TGTGAATACCGCACAGATGCGTAAAGGAGAAATACCGCATCAGCGCTCTTCCGCTTCC	4449	5470	ACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGAGACCCACGC	5529
QY	5268	TGCTCACTGACTCGCTCGCTCGGTTCGGCTGCGCGAGCGGTATCAGTCACTCA	5327	6348	TCACCGGCTCCAGATTTTATCAGCAATAAACACGACCGCGGAGGCGCGAGCAAGT	6407
Db	4450	TGCTCACTGACTCGCTCGCTCGGTTCGGCTGCGCGAGCGGTATCAGTCACTCA	4509	5530	TCACCGGCTCCAGATTTTATCAGCAATAAACACGACCGCGGAGGCGCGAGCAAGT	5589
QY	5328	AAGGCGGTAAATACGTTATCCAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCA	5387	6408	GGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGATA	6467
Db	4510	AAGGCGGTAAATACGTTATCCAGAAATCAGGGGATAACGAGGAAAGAACATGTGAGCA	4569	5590	GGTCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGATA	5649
QY	5388	AAGGCCAGAAAGGCCAGGAACCGTAAAGGCGCGGTGCTGGCGTTTTTCCATAGG	5447	6468	AGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCCATTGCTTACAGGC--TCGTGTG	6526
Db	4570	AAAGGCCAGAAAGGCCAGGAACCGTAAAGGCGCGGTGCTGGCGTTTTTCCATAGG	4629	5650	AGTAGTTCGCGAGTTAATAGTTTGGCAACGTTTGGCCATTGCTTACAGGC--TCGTGTG	5709
QY	5448	CTCCGCCCTCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGGCGAAACCG	5507	6527	TCACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTT	6586
Db	4630	CTCCGCCCTCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGGCGAAACCG	4689	5710	TCACGCTCGTCTGTTGGTATGGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTT	5769
QY	5508	ACAGGACTATAAGATACACAGGCTTTCCCGCTTGAAGCTCCCTCGTGGCTCTCCTGTT	5567	6587	ACATGATCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTGGTCTCCGATCGTTGTC	6646
Db	4690	ACAGGACTATAAGATACACAGGCTTTCCCGCTTGAAGCTCCCTCGTGGCTCTCCTGTT	4749	5770	ACATGATCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTTGGTCTCCGATCGTTGTC	5829
QY	5568	CGACCCCTCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAAGCGTGGCGTT	5627	6647	AGAAGTAGTTGGCGCGAGTGTATCACTCATGTTATGGGAGCACTGCATATTTCTCTT	6706
Db	4750	CGACCCCTCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAAGCGTGGCGTT	4809	5830	AGAAGTAGTTGGCGCGAGTGTATCACTCATGTTATGGGAGCACTGCATATTTCTCTT	5889
QY	5628	TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCCGCTCAGGCTGGC	5687	6707	ACTGTATGCTCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTC	6766
Db	4810	TCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGTCTGCTCCGCTCAGGCTGGC	4869	5890	ACTGTATGCTCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCATTC	5949
QY	5688	TGTGTGACGAACCCCGCTTCAGCCCGGACCGCTGGCGCTTATCCGGTAACTATCGTCTT	5747	6767	TGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATTAATACC	6826
Db	4870	TGTGTGACGAACCCCGCTTCAGCCCGGACCGCTGGCGCTTATCCGGTAACTATCGTCTT	4929	5950	TGAGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGGTCAACACGGGATTAATACC	6009
QY	5748	GAGTCCAAACCGGTAAAGACAGCTTATCGCACTGGGAGAGCCACTGGTAAACGAGATT	5807	6827	GGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGGGAAAA	6886
Db	4930	GAGTCCAAACCGGTAAAGACAGCTTATCGCACTGGGAGAGCCACTGGTAAACGAGATT	4989	6010	GGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAAACGTTCTTCGGGGGAAAA	6069
				6887	CTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCCACTCGTGCACCCAC	6946

Db 6070 CTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACACTCGTGCAACCCCAAC 6129
Qy 6947 TGATCTTCAGCATCTTTTACTTTTACCTTTCACACGCGTTTCTGCGGTGAGCAAAACAGGAAGGCAA 7006
Db 6130 TGAICTTCAGCATCTTTTACTTTTACCTTTCACACGCGTTTCTGCGGTGAGCAAAACAGGAAGGCAA 6189
Qy 7007 AATCGCGCAAAAGGGAATAGGCGGACGCAAGAAATGTTGAATCTCATATCTCTTCCTT 7066
Db 6190 AATCGCGCAAAAGGGAATAGGCGGACGCAAGAAATGTTGAATCTCATATCTCTTCCTT 6249
Qy 7067 TTTCAATATTATGAAGCATTTATCAGGGTATTCTCTCATGAGCGGATACATATTGAA 7126
Db 6250 TTTCAATATTATGAAGCATTTATCAGGGTATTCTCTCATGAGCGGATACATATTGAA 6309
Qy 7127 TGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCACCT 7186
Db 6310 TGTATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCACCT 6369
Qy 7187 GACGCTTAAGAAACCATTTATTCATGACATTAACCTTATAAAATAGCGGTATCACGAGG 7246
Db 6370 GACGCTTAAGAAACCATTTATTCATGACATTAACCTTATAAAATAGCGGTATCACGAGG 6429
Qy 7247 CCCTTTCGCT 7257
Db 6430 CCCTTTCGCT 6440

RESULT 13

AX90482
ID AX90482 standard; DNA; 6522 BP.
XX
AC AX90482;
XX
XX
DT 29-SEP-1999 (first entry)
XX
DE Plasmid retroviral vector pLUSN nucleotide sequence.
XX
KW Plasmid retroviral vector; expression system; immunogenic; gene therapy;
KW immune response; immunosuppression; gene delivery; therapeutic;
KW MHC-I autoimmune disease; tumour; ss.
XX
OS Synthetic.
XX
XX WO936562-A1.
XX
XX
PD 22-JUL-1999.
XX
XX
PF 13-JAN-1999; 99WO-US000733.
XX
XX
PR 14-JAN-1999; 98US-0071409P.
XX
XX (HUMA-) HUMAN GENE THERAPY RES INST.
XX
PI Radosevich TJ, Link CJ;
XX
XX WPI; 1999-468988/39.
XX
XX
PT Expression system containing therapeutic gene and an immunosuppressor
PT gene useful for treating an MHC-I autoimmune disease or killing tumor
PT cells.
XX
XX
PS Claim 25; Page 137-140; 154pp; English.
XX
XX
CC The present invention describes a nucleotide expression system for the
CC introduction of a therapeutic gene comprising: (i) a nucleotide sequence
CC encoding an immune suppression gene; (ii) a promoter; and (iii) a
CC transcription termination signal, where the system is able to inhibit,
CC evade or eliminate a recipient cell immune response to the therapeutic
CC gene when the gene is transformed into a recipient cell. AAX90481 to
CC AAX90484 represent specifically claimed plasmid retroviral vector
CC nucleotide sequences from the present invention. The expression system
CC and vectors containing it can be used for gene therapy, for treating an

CC MHC-I autoimmune disease or for killing tumour cells. The expression
CC system contains an immunosuppressive gene which prevents host rejection
CC of the vector
XX
SQ Sequence 6522 BP; 1469 A; 1792 C; 1704 G; 1557 T; 0 U; 0 Other;
Query Match 40.4%; Score 3096; DB 2; Length 6522;
Best Local Similarity 70.8%; Pred. No. 0;
Matches 4803; Conservative 0; Mismatches 1465; Indels 518; Gaps 31;
Qy 549 TTTTGAAGAGCCACCTGTAGGTTTGCAGCTAGCTTAAAGTAAAGCCATTTTCAAGG 608
Db 174 TTTTGAAGAGCCACCTGTAGGTTTGCAGCTAGCTTAAAGTAAAGCCATTTTCAAGG 231
Qy 609 CATGG-AAAATACATACTGAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAC 667
Db 232 CATGMAAAATACATACTGAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAC 291
Qy 668 AGCAGATATGCGGCCAAACAGGATATCTGTGTAGCGATTCCTGCCCC- GTCAGGGCC 726
Db 292 AGCTGAATA----CCAAACAGGATATCTGTGTAGCGGTTCTGCCCCGCTCAGGGCC 347
Qy 727 AAGAACAGTTGGAACAGGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAGCGATTCCT 785
Db 348 AAGAACAGATGAGACAGCTGAGTGATGGGCCAAACAGGATATCTGTGTAGCGATTCCT 407
Qy 786 GCCCGGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCAGTTTC 845
Db 408 GCCCGGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCAGCCCTCAGCAGTTTC 467
Qy 846 TAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGGACCTG-AAATGACCTGTGCGCTTA 904
Db 468 TAGTGAATCATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAAAATGACCTGTACCTTA 527
Qy 905 TTTGAACTTAACCAATCAGTTGCTTCTCGCTTCTGTGCGCGGTCTTCTGCTCCCGAGCT 964
Db 528 TTTGAACTTAACCAATCAGTTGCTTCTCGCTTCTGTGCGCGGTCTTCTGCTTCCGAGCT 587
Qy 965 CAATAAAGAGCCCAACACCCCTCACTCGGCGCCAGTCTCTCGATAGACTGCGTCGCC 1024
Db 588 CAATAAAGAGCCCAACACCCCTCACTCGGCGCCAGTCTTCCGATAGACTGCGTCGCC 647
Qy 1025 CGGTACCCGTATTTCCCAATAAAGCCTTCTGTGTCATCCGAATCGTGAGCTCGCTG 1084
Db 648 CGGTACCCGTATTTCCCAATAAAGCCTTCTGTGTCATCCGAATCGTGAGCTCGCTG 707
Qy 1085 ATCCTTGGAGGGTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1144
Db 708 TTTCTTGGAGGGTCTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 767
Qy 1145 GGTTCACCGAGATTTGGAGACCCCTGCGGAGGACCAACCGACCCCGCGGAGGTA 1204
Db 768 GGCTCGTCCGGATTTGGAGACCCCTGCGGAGGACCAACCGA-CCACCAACCGGAGGTA 826
Qy 1205 AGCTGGCCAGCACTTATCTGTGTCGTCCGATTTCTAGTGTCTATGACTGATTTATG 1264
Db 827 AGCTGGCCAGCACTTATCTGTGTCGTCCGATTTCTAGTGTCTATGTTGATGTTATG 886
Qy 1265 CGCTCGCTCGGTACTAGTAACTAGCTCTGTATCTGCGGAGCCCGGTGTAAC 1324
Db 887 CGCTCGCTCTGTACTAGTAACTAGCTCTGTATCTGCGGAGCCCGGTGTAAC 946
Qy 1325 GACGAGTTTCGGAAACACCCCGCGCAACCTCGGAGACGTCCTCAGGAGCTTCGGGGCCGT 1384
Db 947 GACGAGTTTCGGAAACACCCCGCGCAACCTCGGAGACGTCCTCAGGAGCTTTGGGGCCGT 1006
Qy 1385 TTTTGTGCGCCGACCTCAGTCTCTAAATCCCGATCGTTTAGGACTCTTTTGTGACACCCC 1444
Db 1007 TTTTGTGCGCCGACCTCAGGAGGAGGAGTCTGATGTAATCCGACCCCGTCA----- 1057
Qy 1445 CTTAGAGAGGAGATATGTGTTCTGTGTTAGGAGAGAGAACCTTAAAGTAGTTCCCGCTCC 1504
Db 1058 -----GGATATGTGTTCTGTGTTAGGAGAGAGAACCTTAAAGTAGTTCCCGCTCC 1107

Db	3245	ATGCCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCGGAATATCATG	3304
Qy	3663	GAGGGGCAAGG--GGCAGCATGGCTTTTACCAGGGTCTCAGTACAGCCACCAAGGACACC	3720
Db	3305	GTGGAAATGGCCGCTTTCTGGATTTCATCGACTGTGGCGGTGGGTGTGGCGGACCGC	3364
Qy	3721	TACGACGCGCTTTCATCATGACGAGCCCTGCGCCCTCGCTTAAC-----TCGACGCGCGCGGGA	3776
Db	3365	TATCAGGACATAGCGTTTGGCTTACCGCTGATATTGCTGAAGAGCTTGGCGGCAATGGGCT	3424
Qy	3777	TCGCGATTAGTCCAATTGTTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGA	3836
Db	3425	GACCGCTTCTGTGCTTTACGGTATCGCGGCTCCGATTCGACGCGCATCGGCTTCTAT	3484
Qy	3837	CAACAATATCACCAAGCTGAAGCCTATAGAGTACGAGCATAGATAAAATAAAAGATTTTA	3896
Db	3485	OGCCTTCTGACGAGTTCTTCTGAGCGGACTCTGGGGTTCGATAAATAAAGATTTTA	3544
Qy	3897	TTTAGTCTCCAGAAAAAGGGGGGAATGAAGACCCACCTGTAGTTTGGTTTGCAGCTAGCT	3956
Db	3545	TTTAGTCTCCAGAAAAAGGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCT	3604
Qy	3957	TAAGTAAGCCATTTTCGACGGCATGG-AAAATACATACTGAGATAGAGAGTTTCAGA	4015
Db	3605	TAAGTAAGCCATTTTCGACGGCATGGAAAAATACATACTGAGATAGAGAGTTTCAGA	3664
Qy	4016	TCAAGGTTTAGGAACAGA-GAGACAGCAGATATATGGGCCAAACAGGATATCTGTGGTAAGC	4074
Db	3665	TCAAGGTCAGGACAGATGGNACAGCTGATATATGGGCCAAACAGGATATCTGTGGTAAGC	3724
Qy	4075	AGTTCCTGCCCC-GCTCAGGCGCCAAACAGTTGGAAACAGGAGATATGGGCCAAACAGG	4133
Db	3725	AGTTCCTGCCCCGCTCAGGCGCCAAACAGATGGAACAGCTGAATATGGGCCAAACAGG	3784
Qy	4134	ATATCTGTGTAGCAGTTCTGTCGCCGCTCAGGGCCAAAGACAGATGTTCCCCAGATG	4193
Db	3785	ATATCTGTGTAGCAGTTCTGTCGCCGCTCAGGGCCAAAGACAGATGTTCCCCAGATG	3844
Qy	4194	CGGTCCCGCCTCAGCAGTTTCTTAGAACAATCAGATGTTTCCAGGGTCCCCCAAGGAC	4253
Db	3845	CGGTCCAGCCTCAGCAGTTTCTTAGAACAATCAGATGTTTCCAGGGTCCCCCAAGGAC	3904
Qy	4254	CTGAATGACCTGTGCTTATTTTGAATAACCAATCAGTTCGCTTCTCGCTTCTGTTCG	4313
Db	3905	CTGAATGACCTGTGCTTATTTTGAATAACCAATCAGTTCGCTTCTCGCTTCTGTTCG	3964
Qy	4314	CGGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGCTCGGCGGGCCAGT	4373
Db	3965	CGGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGCTCGGCGGGCCAGT	4024
Qy	4374	CCCTCGATAGACTGCGTCGCGCGGTTACCGTGTCTCAATAAACCTCTTTCAGTTGCA	4433
Db	4025	CCCTCGATAGACTGAGTCGCGCGGTTACCGTGTATCAATAAACCTCTTTCAGTTGCA	4084
Qy	4434	TCGACTCTGTGCTTCGCTGTTCTTGGGAGGGTCT-CTCTGAGTGATTGACTACCGCTC	4492
Db	4085	TCCGACTTGTGCTCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGATTGACTACCGCTC	4144
Qy	4493	AGCGGGTCTTTCAGTTTCTCCACCTPACAGAGTCTCAATAATTCCTGATGTGCCCGC	4552
Db	4145	AGCGGGGCTTTTCATTTGGGGGCTCTCGCGGATCGGAGACCCCTGCGCCAGGACCA	4204
Qy	4553	AGGAGCTCCGTACGCGCGGTTTTTGTATTATAAATAATGCAAGNACNGTGTTCCTTCAA	4612
Db	4205	CGACCCACACCGGGAGTAAAGTCTG-----	4229
Qy	4613	GCCAGACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGGACTAT	4672
Db	4230	-----GCTGCTTCGGGTTTCGGTGATGACGTGAAACCTCTG-ACACATG	4276
Qy	4673	CTGCCACACGACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATC	4732

Db	4277	CAGCTCCGGAGACGGTCA	CAGCTTGTCTGTAAAGCGATGCCGGAGACGACAAAGCCCGT	4333
Qy	4733	CTCGTCTATTTTGTGGCTTCT	CCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCCTGAG	4792
Db	4337	CAGGGCGGTACAGCGGTGT	TGGCG	4361
Qy	4793	AGCAGGGGATTTGCAGAGT	AGTTAACTCTGAGGGCAGGCTTCTGTGAAAGGTTTCGCT	4852
Db	4362	-----	-----	4361
Qy	4853	GGGCTCAGTGTGAGATTTT	TGCCATAAAAGGGTCTCGCCCTGTGTACAGACAGATCGG	4912
Db	4362	-----	-----	4361
Qy	4913	AATCTAGAGTGCATACT	CAGAGTCCC CGGGTCTCTGATCTCAGGGCATCTTT	4972
Db	4362	-----	-----GGTGTGGGGCGCAGCCATGACCAGTC---	4389
Qy	4973	GCCTAGAGATCCTCTA	CCCGCGACGCATCTGCGCGGGTACCGAGTCTCGAATTCGTAATC	5032
Db	4390	-----	-----ACGTAGCATAGCGGAGTGTATCTATGGCTTTAACTATGC	4427
Qy	5033	ATGGTCTATAGTCTTTT	CTCTGTGTGAATTTGTATCCGCTCACAATTCACACAAACATACG	5092
Db	4428	GGCATCAGAGCAGATTG	TACTGAGA	4452
Qy	5093	AGCCGGAAGCATAAAGT	GTAAGCCTGGGGTGCTAATGATGAGTGAGCTAACTCACATTAAT	5152
Db	4453	-----	-----	4452
Qy	5153	TGCGTTCGGTCACTG	CCGCTTTCCAGTCGGGAAACCTGTCTGTCGACGCTGCATTAATG	5212
Db	4453	-----	-----GTGCACCATATGCGGTGTGA	4472
Qy	5213	AATCGGCCAACGCGCGGG	AGAGCGGTTTGCGTATTTGGCGCTCTTCCGCTTCTTCGCT	5272
Db	4473	AATACCCGACAGAT	TGCGTAAGAGGAAATACCCGATCAGGCGCTCTTCCGCTTCTTCGCT	4532
Qy	5273	CACGTACTCGCTCGCT	CGCTGTGGCTGCGCGAGCGGTATCAGTCACTCAAAAGGC	5332
Db	4533	CACGTACTCGCTGGCT	CGGTCTGGCTGCGCGAGCGGTATCAGTCACTCAAAAGGC	4592
Qy	5333	GGTAATACGGTTATCCA	CAGAATCAGGGGATAACGAGGAAAGAAATGTGAGCAAAAGG	5392
Db	4593	GGTAATACGGTTATCCA	CAGAAATCAGGGGATAACGAGGAAAGAAATGTGAGCAAAAGG	4652
Qy	5393	CCAGCAAAAGGCCAGG	AAACGTAAGGCGCGTGTCTGGCGCTTTTTCATAGGCTCCG	5452
Db	4653	CCAGCAAAAGGCCAGG	AAACGTAAGGCGCGTGTCTGGCGCTTTTTCATAGGCTCCG	4712
Qy	5453	CCCCCTGACGAGCAT	CACAAAAATCGACGCTCAAGTCTGAGGTCGCGAAACCCGACAGG	5512
Db	4713	CCCCCTGACGAGCAT	CACAAAAATCGACGCTCAAGTCTGAGGTCGCGAAACCCGACAGG	4772
Qy	5513	ACTATAAGATACAG	CGGCTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTTCCGAC	5572
Db	4773	ACTATAAGATACAG	CGGCTTTCCCTCGTGAAGCTCCCTCGTGGCTCTCTCTTCCGAC	4832
Qy	5573	CTTGGCGCTTACCG	GATACCTGTCCGCTTTTCTCTTCTGGGAAGCGTGGCGTTTCTCA	5632
Db	4833	CCTGGCGCTTACCG	GATACCTGTCCGCTTTTCTCTCTTCTGGGAAGCGTGGCGTTTCTCA	4892
Qy	5633	TAGCTACGCTGTAG	GTATCTCAGTTCTGGGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGT	5692
Db	4893	TAGCTACGCTGTAG	GTATCTCAGTTCTGGGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGT	4952
Qy	5693	GCAGAACCCCGCTT	CAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTC	5752
Db	4953	GCAGAACCCCGCTT	CAGCCGACCGCTGCGCTTATCCGGTAACTATCGTCTTGAGTC	5012
Qy	5753	CAACCCGGTAAACA	CGACTTATCGCACTGCGAGCGACCTGGTAAACAGGATTAAGCAG	5812
Db	5013	CAACCCGGTAAACA	CGACTTATCGCACTGCGAGCGACCTGGTAAACAGGATTAAGCAG	5072


```
QY 5813 AGCGAGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACAC 5872
Db |||||||
QY 5073 AGCGAGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACAC 5132
Db |||||||
QY 5873 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACCTTCGGAAAAAGAGT 5932
Db |||||||
QY 5133 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACCTTCGGAAAAAGAGT 5192
QY 5933 TGGTAGCTCTTGATCCGCGCAAAACAACACCGCTGGTAGCGGTGGTTTTTTTGTGCAAA 5992
Db |||||||
QY 5193 TGGTAGCTCTTGATCCGCGCAAAACAACACCGCTGGTAGCGGTGGTTTTTTTGTGCAAA 5252
QY 5993 GCAGCAGATTAGCGCGAGAAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTCTACGGG 6052
Db |||||||
QY 5253 GCAGCAGATTAGCGCGAGAAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTTCTACGGG 5312
QY 6053 GTCTGACGCTCAGTGGAACGAAACCTCACGTTAAGGGATTTTGGTCATGAGATTAACAAA 6112
Db |||||||
QY 5313 GTCTGACGCTCAGTGGAACGAAACCTCACGTTAAGGGATTTTGGTCATGAGATTAACAAA 5372
QY 6113 AAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6172
Db |||||||
QY 5373 AAGGATCTTCACTAGATCCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 5432
QY 6173 ATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAGGCACCTATCTCAGC 6232
Db |||||||
QY 5433 ATATGAGTAAACTTGGTCTGACAGTTTACCAATGCTTTAATCAGTGAGGCACCTATCTCAGC 5492
QY 6233 GATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGAT 6292
Db |||||||
QY 5493 GATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGAT 5552
QY 6293 ACGGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGCGAGACCGAGCTCACC 6352
Db |||||||
QY 5553 ACGGGAGGCTTACCATCTGGCCCGAGTCTGCAATGATACCGCGAGACCGAGCTCACC 5612
QY 6353 GGCTCCAGATTTATCAGCAATAAACCAGCAGCGCGGAGCGCGAGCAAGTGGTCC 6412
Db |||||||
QY 5613 GGCTCCAGATTTATCAGCAATAAACCAGCAGCGCGGAGCGCGAGCAAGTGGTCC 5672
QY 6413 TGCAACTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAG 6472
Db |||||||
QY 5673 TGCAACTTTATCGGCTCCATCCAGTCTAATTAATTTGTCGGGAGCTAGAGTAAGTAG 5732
QY 6473 TTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGC-TGCTGGTGTACAG 6531
Db |||||||
QY 5733 TTGCGCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTGAGGCATCGTGGTGTACAG 5792
QY 6532 CTCGTCGTTTGGTATGGCTTTCAATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATG 6591
Db |||||||
QY 5793 CTCGTCGTTTGGTATGGCTTTCAATTCAGCTCCGTTCCCAACGATCAAGGCGAGTTACATG 5852
QY 6592 ATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTGTCAAGAG 6651
Db |||||||
QY 5853 ATCCCCCATGTTGTGCAAAAAAGCGGTAGTCTCTTCGGTCTCCGATCGTTGTCAAGAG 5912
QY 6652 TAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATCAATTTCTTTACTGT 6711
Db |||||||
QY 5913 TAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATCAATTTCTTTACTGT 5972
QY 6712 CATGCCATCCGTAAGATGTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTTGAGA 6771
Db |||||||
QY 5973 CATGCCATCCGTAAGATGTTTTCTGTGACTGTTGAGTACTCAACCAAGTCAATCTTGAGA 6032
QY 6772 ATAGTGTATCGCGCGACCGAGTTGCTCTTTCGCCGGCGTCAATAACGGGATAATACCGCGCC 6831
Db |||||||
QY 6033 ATAGTGTATCGCGCGACCGAGTTGCTCTTTCGCCGGCGTCAACACGGGATAATACCGCGCC 6092
QY 6832 ACATAGCAGAACTTTTAAAGTGTCTCATTTGGGAAACGTTCTTCGGGGCGGAAACTCTC 6891
Db |||||||
QY 6093 ACATAGCAGAACTTTTAAAGTGTCTCATTTGGGAAACGTTCTTCGGGGCGGAAACTCTC 6152
```

```
QY 6892 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCAACTGATC 6951
Db |||||||
QY 6153 AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGCAACCAACTGATC 6212
QY 6952 TTACAGATCTTTTACTTTTACCGAGGTTTCTGGGTGAGCAAAAACAGGAGGCAAAATGC 7011
Db |||||||
QY 6213 TTACAGATCTTTTACTTTTACCGAGGTTTCTGGGTGAGCAAAAACAGGAGGCAAAATGC 6272
QY 7012 CCAAAAAAGGGAATAAAGGGGACACGGAATGTTGAATACTCATACTCTCTCTTTTCA 7071
Db |||||||
QY 6273 CCAAAAAAGGGAATAAAGGGGACACGGAATGTTGAATACTCATACTCTCTCTTTTCA 6332
QY 7072 ATATTATTGAAGCAATTTATCAGGGTTATTGTCATGAGCGGATACATATTTGAATGAT 7131
Db |||||||
QY 6333 ATATTATTGAAGCAATTTATCAGGGTTATTGTCATGAGCGGATACATATTTGAATGAT 6392
QY 7132 TTAGAAAAATAAACAATAAGGGTTCCGGGCACATTTCCCGGAAAAAGTGCACCTGACGT 7191
Db |||||||
QY 6393 TTAGAAAAATAAACAATAAGGGTTCCGGGCACATTTCCCGGAAAAAGTGCACCTGACGT 6452
QY 7192 CTAAGAAACCATTATTATCATGACATTAACCTTATAAAATAGGCGGTATCAGGAGGCCCTT 7251
Db |||||||
QY 6453 CTAAGAAACCATTATTATCATGACATTAACCTTATAAAATAGGCGGTATCAGGAGGCCCTT 6512
QY 7252 TCGTCT 7257
Db |||||||
QY 6513 TCGTCT 6518
Db |||||||

RESULT 14
AAZ34935
ID AAZ34935 standard; DNA; 6221 BP.
XX
AC AAZ34935;
XX
DT 28-FEB-2000 (first entry)
XX
DE Retrovirus vector.
XX
KW Interleukin-4 inducible epsilon promoter; immunoglobulin; IgE; antibody;
KW modulator; screening; human; allergy; therapy; retrovirus; vector;
KW green fluorescent protein; GFP; ss.
XX
OS Cytomegalovirus.
OS Aequorea victoria.
OS Unidentified.
OS Chimeric.
XX
FH Key
FT promoter
FT 1. .845
FT /tag= a
FT /note= "CMV promoter/R/US 5' LTR"
FT
FT misc_feature
FT 850. .2100
FT /tag= c
FT /note= "extended psi region"
FT
FT mutation
FT replace(1322,G)
FT /tag= b
FT /note= "CAG ATG-ATC mutation"
FT
FT misc_feature
FT 2205. .2723
FT /tag= d
FT /note= "ECMV IRES cloned from pCITE-4a"
FT
FT CDS
FT 2746. .3465
FT /tag= e
FT /product= "green fluorescent protein"
FT
FT LTR
FT 3522. .4115
FT /tag= f
FT /note= "3' LTR"
FT
FT misc_feature
FT 4122. .62100
FT /tag= g
FT /note= "pGEM backbone (pUC origin, ampR)"
XX
XX WO9958663-A1.
XX
```


Db 2203 CTAGAGCGAATTAATTCGGTATTTTCCACCA-----TA 2237
Qy 2576 CTTCTGTAGCTTGTTACCAAGCAAGCCAGGTAAAGCTCCAAAGCTGTGATCTACTGGA 2635
Db 2238 TTGCGCTCTTTTGGCNAATGAGGGCCGGAACCTGGGCCCTGTCTTTTGACGACATT 2297
Qy 2636 CATCCACCGCGCACACTGGTGTGCCAAGCAGATTCAGCGGTAGCGGTACCGACT 2695
Db 2298 CTTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGAATGTCGTGAAGAA 2357
Qy 2696 TCACCTTCAACATCAGACGCTCCAGCCAGAGACATCGCCACCTACTACTCTGCGACAA 2755
Db 2358 GCAGTTCTCTGGAAGCTTTCTGAAGACAAACAACTCTGTAGCACCTT----- 2408
Qy 2756 ATAGCTCTATCGGTGTTGCGCAAGGACCAAGGTGGAATCAACAGAGTGGCTCAG 2815
Db 2409 -----TGAGGCGAGCGAAACCCCCCACTGGCGACAGTGCCTTGGC 2451
Qy 2816 GATCGGGTGGATCCGGCTCTGGTGGCTCAGGATCGAGGCTCAACTGTGGAGAGCGGTG 2875
Db 2452 GCCAAAGCCACGTGTATAGATACACTGCAAGGGGSCAACAACCCAGTGCACGTTG 2511
Qy 2876 GAGGTGTTGCAACCTGCGCGGTCCCTGGGCTGCTGCTGCTCCGATCTGGCTTCGATT 2935
Db 2512 TGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAGGGGC 2571
Qy 2936 TCACCACATATTTGGATAGTTGGGTGAGACAGCACCTGGAAAAGTCTTTGATGGATTG 2995
Db 2572 TGAAGGATCCCAAGAGGTACCCCATTTGTAGGATCTGATCTGGGGCTCGGTGCACAT 2631
Qy 2996 GAGAAATTCATCAGATAGCAGTACGATTAACTATGCGCGCTCTTAAAGGATAGATTTA 3055
Db 2632 GCTTTACATGTGTTAGTCGAGTTAAAAAAGTCTAGGCCCCCGAACCAACCGGACGT 2691
Qy 3056 CAATATCGGAGACAAAGCCAGAAACACATTTGCTGCAAAATGAGACGCTGAGACCGG 3115
Db 2692 GGTGTTTCTTTGAAAAACAGATGATAATG-----GGGGATCCACCGG 2736
Qy 3116 AAGACACCGGGTCTATTTTGTGNAAGCTTTACTTCGGCTTCCCTGGTTGCTTATT 3175
Db 2737 TCGCCACCATGTGTAGCAAGGGCGAGAGCTGTTCACCGGGTGG--TGCCCATCTGGT 2794
Qy 3176 GGGGCAAGGGACCCCGGTACCGTCTCCAGTGTAAAGCCACACGAGCCAGCGCCGCG 3235
Db 2795 CGAGCTGGACGGCAGGTAAACGGCCACAAGTTACGCTGTCCGGCAGGGCGAGGGCA 2854
Qy 3236 GACCACCAACACCGGCGCCCAATCGGTGCGAGCCCTGTCTGCGCCCGCAGAGGGCG 3295
Db 2855 TGCACCTACGGCA-AGCTGACCTGAAGTTCACTGCAACACCGGCAAGCTGCCGTGC 2913
Qy 3296 CTGCGGCAGGGGGGGGGCGAGTGCAACAGAGGGGTGAGACTTCGGCCCTGGATCCCA 3355
Db 2914 CTTGGCCCACTCTGAGACACCTGACCTAGCCGCTAGCGGTGAGTGTCTCAGCCGTACCCCG 2973
Qy 3356 AACTCTGCTACCTGTGTTGGAATGGAATCCTCTTCACTATGTTGTTCTCACTGCTTGT 3415
Db 2974 ACCACATGAGACGACGA-----CTTCTTAAGTCCGCCATGCCCAGAGGTACG 3024
Qy 3416 TCCGTAGAGTGAAGTTACAGAGAGCGAGAGCCCGCGGTACCAAGAGGGCGCAAGAAC 3475
Db 3025 TCCAGGAGCGCACCATCTTTCTTCAAGACGAGGCAACTACAAGACCCGCGCGAGGTGA 3084
Qy 3476 AGCTCTATTAACGAGCTCAATCTAGGACGAAGAGAGTACGATGTTTTGGACAGAGAC 3535
Db 3085 AGTTCAGGGCGGACACCTCGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG 3144
Qy 3536 GTGGCCGGGACCTGAGATGGGGGAAAGCCGAGAGGAGAACCTCAGG--AAAGGCC 3592
Db 3145 ACGGCAACATCTGGGGGCAAGCTGGAGTACAACTACACAGCCCAACAGCTTATATCA 3204
Qy 3593 TGTACAAATGACATGCAAGAAGATAGATGGCGAGGCCCTACAGTGAATTTGGGATGAAAG 3652

Db 3205 TGGCCGCAAGCAGAGAAGAACGGCATCAAGGTGAATCTCAAGATCCGCCACAACATCGAGG 3264
Qy 3653 GCGAGCGCGGAGGGGCAAGGGCACGATGGGCTTTTACAGGGTCTCAGTACAGCACCA 3712
Db 3265 ACGGCAGCGTGCAGTCTGCCGACCACTACACAGCAGAACACCCCCATCGGCGAGCGCCCG 3324
Qy 3713 AGGACACCTTACGACGCCCTTACATGACAGGCCCTCGCCCTCGCTAACTCTGACGGCGCG 3772
Db 3325 TGCTGCTGCCCAACCACTACCTGAGCAGCCAGCTCGGCCCTGAGCAAGAACCCCAACG 3384
Qy 3773 CGGATCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 3832
Db 3385 AGAAGCGCGATCA-----CATGCTCTCTGTGGAGTTTCGTGACCGCCCGCGGAT 3433
Qy 3833 GACTCAACAATATCACAGCTGAAGCTATAGATACGAGCCATAGATAAATAAATAAAGAT 3892
Db 3434 CACTCTCGCATGAGAGAGCTGTACAAGTAAGCGCGCTCGACGATAAATAAAGAT 3493
Qy 3893 TTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGGTTTGGCAAGCT 3952
Db 3494 TTTATTTAGTCTCCAGAAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGGCAAGCT 3553
Qy 3953 AGCTTAAGTAAGCCATTTTGAAGGCATGG--AAATACATAACTGAGNATAGAGAGTT 4011
Db 3554 AGCTTAAGTAAGCCATTTTGAAGGCATGGAAAAATACATAACTGAGNATAGAGAGTT 3613
Qy 4012 CAGATCAAGTTTGAAGAACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGT 4070
Db 3614 CAGATCAAGTCTCAGAACAGATGGAAACAGCTGAATATGGGCCAAACAGGATATCTGTGT 3673
Qy 4071 AAGCAGTTCTCTGCCCC-GTCAAGGCCAAGAACAGTTGGAAACAGAGAAATATGGGCCAAA 4129
Db 3674 AAGCAGTTCTCTGCCCCGCTCAGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 3733
Qy 4130 CAGGATATCTGTGTTAAGAGTTCTGCCCGGCTCAGGCCCAAGAACAGATGTTGCCCA 4189
Db 3734 CAGGATATCTGTGTTAAGCAGTTCTGCCCGGCTCAGGCCCAAGAACAGATGTTGCCCA 3793
Qy 4190 GATCGGTCCCGCTCAGCAGTTCTTAGAAGAACCATCAGATGTTTTCAGGGGTGCCCA 4249
Db 3794 GATCGGTCCAGCCCTCAGCAGTTCTTAGAAGAACCATCAGATGTTTTCAGGGGTGCCCA 3853
Qy 4250 GGACCTGAAATGACCTGTGCTTATTTGAATCAACCAATCAGTTTCTCGTTCTG 4309
Db 3854 GGACCTGAAATGACCTGTGCTTATTTGAATCAACCAATCAGTTTCTCGTTCTG 3913
Qy 4310 TTCGCGCGTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCG 4369
Db 3914 TTCGCGCGTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCG 3973
Qy 4370 CAGTCTCCGATAGACTCGTCCCGGGTACCCGTTTCTCAATAAACCCCTCTTGCAGT 4429
Db 3974 CAGTCTCCGATAGACTGAGTCCCGGGTACCCGTTTATCCAAATAAACCCCTCTTGCAGT 4033
Qy 4430 TGCATCCGACTCGTGTCTGCTGTTTCTTGGAGGGTCT-CTCTGAGTGAATGACTACC 4488
Db 4034 TGCATCCGACTGTTGTCTGCTGTTTCTTGGAGGGTCTCTCTGAGTGAATGACTACC 4093
Qy 4489 CQTACGCGGGTCTTTCAGTTTCTCCCACTACAGGTTCTCACTAACTTCTGATG 4548
Db 4094 CQTACGCGGGTCTTTCATT----- 4114
Qy 4549 CCGCAGGACTCCGTGAGCCCGTTTGTGTTTATAAATAAAGAACAGTGTTCCT 4608
Db 4115 ----- 4114
Qy 4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGATGTTGAAGGCTCTGTGA 4668
Db 4115 ----- 4114
Qy 4669 CTATCTGCCACAGACTTTTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATC 4728
Db 4115 ----- 4114

QY 4729 TATCCTCGTCTATTTTGGCTTCTCCGTATTTTAAATTTCTAGTTTGGCACTCCCTTCC 4788
Db 4115 ----- 4114
QY 4789 TGAGACACGGCGATTGCGAGAGTAGTAACTACTGTAGGCGAGGCTTCTGTGAAAAGGTT 4848
Db 4115 -----TCGACTGTGGTCTCGTGCCTTGGGA 4142
QY 4849 GCCTGGGCTCAGTGTGAGATTGTCATAAAAAGGGTCTCGCCCTCTGTACAGACAGA 4908
Db 4143 GGGTCTCTCTGAGTGATTGACTACCGCTCAGCGGGGTCTTCATGC----- 4191
QY 4909 TCGGAATCTAGAGTCATACTCAGAGTCCC CGCGGTTCCGGGGTCTCTGATCTCAGGGCAT 4968
Db 4192 ----- 4191
QY 4969 CTTTGCCTAGAGATCCTCTACGCCGAGCATCGTGGCGGGTACCGAGCTCGAATTGCT 5028
Db 4192 ----- 4191
QY 5029 AATCATGTCTATGCTGTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCACACACA 5088
Db 4192 -----AGCATGTACAAATTAATTTGGTTTTTTTCTTAAGTATTACATTA 4239
QY 5089 TACGAGCCGGAAGCATAAAGTGAAGCTCGGGTGCCTAATGATGAGCTAACTCAAT 5148
Db 4240 AA----- 4241
QY 5149 TAAITGGCTTGGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTGTGCGCAGCTGCATT 5208
Db 4242 -----TGCCCATAGTTGCAATT 4257
QY 5209 AATGAATCGGCAACCGCGGGAGAGCGGTTTTCGTATTTGGCGCTCTTCCGCTTCT 5268
Db 4258 AATGAATCGGCAACCGCGGGAGAGCGGTTTTCGTATTTGGCGCTCTTCCGCTTCT 4316
QY 5269 CGCTCACTGA CTGCTGCTCGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 5328
Db 4317 CGCTCACTGA CTGCTGCTCGTTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 4376
QY 5329 AGCGGTAATACGGTTATCCACAGATCAGGGGATACGAGGAAAGACATGTAGCAA 5388
Db 4377 AGCGGTAATACGGTTATCCACAGATCAGGGGATACGAGGAAAGACATGTAGCAA 4436
QY 5389 AAGGCCAGAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTTTTCCATAGGC 5448
Db 4437 AAGGCCAGAAAAGGCCAGGAACCGTAAAGGCGCGTTCGTCGCGTTTTTCCATAGGC 4496
QY 5449 TCCGCCCGCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 5508
Db 4497 TCCGCCCGCTTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 4556
QY 5509 CAGGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTTC 5568
Db 4557 CAGGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTGTTTC 4616
QY 5569 CGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTT 5628
Db 4617 CGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTT 4676
QY 5629 CTCATAGTCACTGAGTATCTCAGTTTCGGTGTAGGTGTTTCGCTCCAAAGCTGGGCT 5688
Db 4677 CTCATAGTCACTGAGTATCTCAGTTTCGGTGTAGGTGTTTCGCTCCAAAGCTGGGCT 4736
QY 5689 GTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCTGTTG 5748
Db 4737 GTGTGACGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCTGTTG 4796
QY 5749 AGTCCAAACCCGGTAAGACACGACTTATCGCCACTTGGCAGGACGCTGTTAAACAGGATTA 5808
Db 4797 AGTCCAAACCCGGTAAGACACGACTTATCGCCACTTGGCAGGACGCTGTTAAACAGGATTA 4856

QY 5809 GCAGAGCGAGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTTAACCTACGCT 5868
Db 4857 GCAGAGCGAGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTTAACCTACGCT 4916
QY 5869 ACACCTAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGAAAAA 5928
Db 4917 ACACCTAGAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGAAAAA 4976
QY 5929 GAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGGGTGGTTTTTTTGT 5988
Db 4977 GAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGTGTAGGGTGGTTTTTTTGT 5036
QY 5989 GCAAGCAGCAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCTTTTGTATCTTA 6048
Db 5037 GCAAGCAGCAGATTACGCGCAGAAAAAAGAGATCTCAAGAAAGATCTTTTGTATCTTA 5096
QY 6049 CGGGTCTGAGCTCAGTGGAAACCACTCAGCTTAAGGATTTTGGTCAATGAGATTAT 6108
Db 5097 CGGGTCTGAGCTCAGTGGAAACCACTCAGCTTAAGGATTTTGGTCAATGAGATTAT 5156
QY 6109 CAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTT---TAAATCAATCT 6165
Db 5157 CAAAAAGGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTGGCGCAATCAATCT 5216
QY 6166 AAGTATATATGAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTA 6225
Db 5217 AAGTATATATGAGTAACTTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGCACCTA 5276
QY 6226 TCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 6285
Db 5277 TCTCAGCGATCTGCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTCTGTAGATAA 5336
QY 6286 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAC 6345
Db 5337 CTACGATACGGGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGCGAGACCCAC 5396
QY 6346 GCTCACCGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAGGCGCAGCAGAA 6405
Db 5397 GCTCACCGCTCCAGATTTATCAGCAATAAACAGCCAGCGGAGGCGCAGCAGAA 5456
QY 6406 GTGCTCTGCAACTTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAG 6465
Db 5457 GTGCTCTGCAACTTTTATCCGCTTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAG 5516
QY 6466 TAACTAGTTCGCGAGTTAATAGTTTGCACAACTGTTGCCATTTGCTACAGGC---TCGTGG 6524
Db 5517 TAACTAGTTCGCGAGTTAATAGTTTGCACAACTGTTGCCATTTGCTACAGGCATCTGTCG 5576
QY 6525 TGTCAAGCTCTGCTGTTGGTATGCTTTCACTCAGCTCCGTTCCCAACGATCAAGGCGAG 6584
Db 5577 TGTCAAGCTCTGCTGTTGGTATGCTTTCACTCAGCTCCGTTCCCAACGATCAAGGCGAG 5636
QY 6585 TTACATGATCCCCATGTTGTGCAAAAAGGGTGTAGCTCTTCGCTCTCCGATCGTTG 6644
Db 5637 TTACATGATCCCCATGTTGTGCAAAAAGGGTGTAGCTCTTCGCTCTCCGATCGTTG 5696
QY 6645 TCAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGATTAATCTC 6704
Db 5697 TCAGAGTAAAGTTGGCGCAGTGTATCACTCATGTTTATGGCAGCACTGATTAATCTC 5756
QY 6705 TTAATGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAT 6764
Db 5757 TTAATGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAT 5816
QY 6765 TCTGAGATAGTGTATGCGGGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATATA 6824
Db 5817 TCTGAGATAGTGTATGCGGGGACCGAGTTGCTCTTCCCGCGGTCAATACGGGATATA 5876
QY 6825 CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATTGGAACAGTTCTTCGGGGCGAA 6884
Db 5877 CCGCGCACATAGCAGAACTTTTAAAGTGTCTCATATTGGAACAGTTCTTCGGGGCGAA 5936
QY 6885 AACTCTCAAGGATCTTACCCTGTTTGGATCCAGTTTCGATGTAAACCACTCGTGCACCCA 6944

Db 5937 AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACCTGCTGCACCA 5996
Qy 6945 ACTGATCTTCAGCATCTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGC 7004
Db 5997 ACTGATCTTCAGCATCTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGC 6056
Qy 7005 AAAATGCCCAAAAAGGAAATAAGGGCGACACGGAATGTTGATACTCATACTCTTCC 7064
Db 6057 AAAATGCCCAAAAAGGAAATAAGGGCGACACGGAATGTTGATACTCATACTCTTCC 6116
Qy 7065 TTTTTCATATTATTGAAGCATTTATFACGGGTTATTGTCTCATGAGCGGATACATATTG 7124
Db 6117 TTTTTCATATTATTGAAGCATTTATFACGGGTTATTGTCTCATGAGCGGATACATATTG 6176
Qy 7125 AATGATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTTC 7169
Db 6177 AATGATTTAGAAAAATAAACAATAGGGGTTCCGGCGACATTTTC 6221

RESULT 15

AAF30944
ID AAF30944 standard; DNA; 6221 BP.

AC AAF30944;

DT 11-SEP-2003 (revised)

DT 23-JUL-2001 (first entry)

XX Vector used in invention.

XX Interleukin-4 inducible epsilon promoter; human; IgG; antibody;

KW immunoglobulin E; allergy; therapy; switch rearrangement; vector; CMV;

KW green fluorescent protein; ds.

OS Cytoomegalovirus.

OS unidentified.

OS Chimeric.

FH Key Location/Qualifiers

FT LTR 1..845

FT /tag= a

FT /note= "5' LTR"

FT promoter 1..845

FT /tag= b

FT /note= "CMV promoter"

FT misc_feature 850..2100

FT /tag= d

FT /note= "extended psi region"

FT mutation replace(1322,G)

FT /tag= c

FT misc_feature 2146..2173

FT /tag= e

FT /note= "2 Bstx1 peptide cloning sites"

FT misc_feature 2205..2723

FT /tag= f

FT /note= "ECMV IRES cloned as EcoRI/MscI fragment from

FT pCITE-4a"

FT CDS 2745..3464

FT /tag= g

FT /product= "green fluorescent protein"

FT LTR 3522..4115

FT /tag= h

FT /note= "3' LTR"

FT misc_feature 4122..6210

FT /tag= i

FT /note= "pGEM backbone (pUC origin, ampr)"

XX WO200134806-A2.

XX 17-MAY-2001.

XX 13-NOV-2000; 2000WO-US031232.

PF

XX 12-NOV-1999; 99US-0165189P.
PR (RIGE-) RIGEL PHARM INC.
PA Kinsella TW;
XX WPI; 2001-335931/35.
XX Screening for agents capable of inhibiting a promoter, especially
PT interleukin-4 inducible epsilon promoter involved in immunoglobulin E
PT production, by using diphtheria toxin constructs.
XX Disclosure; Fig 11A-1-11A-3; 80pp; English.
XX The present sequence is that of a vector preferred for use in methods of
CC the invention. These methods utilise diphtheria toxin for screening
CC purposes, especially for identifying modulators of IgE synthesis,
CC secretion and switch rearrangement. A claimed method of screening for
CC bioactive agents capable of inhibiting the IL-4 inducible epsilon
CC promoter, which is involved in IgE switching, comprises: combining a
CC candidate bioactive agent and a cell that does not endogenously express
CC heparin-binding epidermal growth factor-like growth factor (HBEGF) and
CC which comprises a fusion nucleic acid encoding HBEGF; inducing the promoter
CC epsilon promoter and a nucleic acid encoding HBEGF; inducing the promoter
CC with IL-4; adding diphtheria toxin to the cell; and determining whether
CC the cell is dead. Compositions comprising a test vector and a reporter
CC vector that includes a reporter gene such as green fluorescent proteins
CC are provided. Inhibitors of IgE synthesis can be identified that prevent
CC the production of IgE and reduce or eliminate an allergic response.
CC (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 6221 BP; 1471 A; 1738 C; 1548 G; 1464 T; 0 U; 0 Other;
Query Match 40.4%; Score 3093; DB 4; Length 6221;
Best Local Similarity 73.7%; Pred. No. 0;
Matches 4587; Conservative 0; Mismatches 960; Indels 678; Gaps 24;
Qy 958 CCGAGCTCAATAAAGAGCCACACCCCTCACTCGGGCGCCAGTCTCCCATGACATG 1017
Db 662 CAGAGCTCAATAAAGAGCCACACCCCTCACTCGGGCGCCAGTCTCCCATGACATG 721
Qy 1018 CGTCGCCGGGTACCCGTATTCCCAATAAAGCTCTTGTGTTCATCCGAATCGTGGA 1077
Db 722 AGTCGCCGGGTACCCGTATTCCCAATAAAGCTCTTGTGTTCATCCGAATCGTGGA 780
Qy 1078 CTCGCTGATCTCTGGGAGGGTCTCTCAGATTGATGATGATGCC-ACCTCGGGGGTCTTT 1136
Db 781 CTCGCTGTTCTCTGGGAGGGTCTCTCAGATTGATGATGATGCC-ACCTCGGGGGTCTTT 840
Qy 1137 CATTTGGAGGTTCCACCGAGATTTGGAGACCCCTCCAGGAGCACCAGACCCCGCCG 1196
Db 841 CATTTGGAGGTTCCACCGAGATTTGGAGACCCCTCCAGGAGCACCAGACCCCGCCG 899
Qy 1197 GGGAGGTAAAGCTGGCCAGCACTTATCTGTCTGTCTGCGATTGTCTAGTGTCTATGAC 1256
Db 900 GGGAGGTAAAGCTGGCCAGCACTTATCTGTCTGTCTGCGATTGTCTAGTGTCTATGAC 959
Qy 1257 ATTTATGCGCTGCTGGTACTAGTACTAGTAACTAGCTCTGTATCTGCGGACCCGCTG 1316
Db 960 ATTTATGCGCTGCTGGTACTAGTACTAGTAACTAGCTCTGTATCTGCGGACCCGCTG 1019
Qy 1317 GTGGAACCTGACGATTCCGACACCCGCCCAACCCCTGGGAGAGCTCCACGGAGCTTCG 1376
Db 1020 GTGGAACCTGACGATTCCGACACCCGCCCAACCCCTGGGAGAGCTCCACGGAGCTTCG 1079
Qy 1377 GGGGCGGTTTTGTGGCCCGGACCTGAGTCTCTAAATCCCGATCGTTTAGACTCTTTGGT 1436
Db 1080 GGGGCGGTTTTGTGGCCCGGACCTGAGTCTCTAAATCCCGATCGTTTAGACTCTTTGGT 1139
Qy 1437 GCACCCCTTTAGAGAGGATATGTGGTCTGTGGTAGGAGACGAGAACCTAAACAGTTTC 1496
Db 1140 GCACCCCTTTAGAGAGGATATGTGGTCTGTGGTAGGAGACGAGAACCTAAACAGTTTC 1199

QY 1497 CCGCCTCGGTCTGAAATTTTGGTTTGGTTTGGGACCGAAGCCGCGCGCGGTCTTGTCT 1556
DB |||||
1200 CCGCCTCGGTCTGAAATTTTGGTTTGGTTTGGGACCGAAGCCGCGCGCGGTCTTGTCT 1259
QY |||||
1557 TGCTCAGCATCGTTCTGTTGTTGTTCTCTGCTCTGATCTGTTTCTGTTTCTGTTTCTGTTT 1616
DB |||||
1260 TGCTCAGCATCGTTCTGTTGTTGTTCTCTGCTCTGATCTGTTTCTGTTTCTGTTTCTGTTT 1319
QY |||||
1617 ATGGCCCGGGGTAGACTGTACCACTCCCTTTAAGTTTGGCTTTAGGTCACCTGGAAAGAT 1676
DB |||||
1320 ATGGCCCGGGGTAGACTGTACCACTCCCTTTAAGTTTGGCTTTAGGTCACCTGGAAAGAT 1379
QY |||||
1677 GTGAGCGGATCGCTCACAACAGTCTGGTAGATGTCAAGAGAGAGCTTGGTTTACCTTC 1736
DB |||||
1380 GTGAGCGGATCGCTCACAACAGTCTGGTAGATGTCAAGAGAGAGCTTGGTTTACCTTC 1439
QY |||||
1737 TGCTCTGAGATGCGCAACCTTTAAGTCTGGATGCGCGGAGCGGACCTTTAAGCGA 1796
DB |||||
1440 TGCTCTGAGATGCGCAACCTTTAAGTCTGGATGCGCGGAGCGGACCTTTAAGCGA 1499
QY |||||
1797 GACCTCATCACCGGTTAAGATCAAGGTCTTTTCACTGGCGCGCATGGACACCCAGAC 1856
DB |||||
1500 GACCTCATCACCGGTTAAGATCAAGGTCTTTTCACTGGCGCGCATGGACACCCAGAC 1559
QY |||||
1857 CAGTCCCTCATCATCGTGAACCTGGAAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAG 1916
DB |||||
1560 CAGTCCCTCATCATCGTGAACCTGGAAAGCTTGGCTTTTGAACCCCTCCCTGGGTCAAG 1619
QY |||||
1917 CCGTTTGTACACCTTAAGCTCCGCTCTCTTCTCTCCATCGCGCGCGCTCTCCCTTT 1976
DB |||||
1620 CCGTTTGTACACCTTAAGCTCCGCTCTCTTCTCTCCATCGCGCGCGCTCTCCCTTT 1679
QY |||||
1977 GAACCTCTCGTTTCGACCGCGCTCGATCTCCCTTTTATCAAGCCCTCACTCTTCTCTA 2036
DB |||||
1680 GAACCTCTCGTTTCGACCGCGCTCGATCTCCCTTTTATCAAGCCCTCACTCTTCTCTA 1739
QY |||||
2037 GCGCGCCCATATGCGCATATGAGATCTTATATGGGCAACCCCGCTTGTAACTTC 2096
DB |||||
1740 GCGCGCCCATATGCGCATATGAGATCTTATATGGGCAACCCCGCTTGTAACTTC 1799
QY |||||
2097 CCGTGAACCTGATCAGAGAGTACTTAACAGCCCTCTCTCAAGCTCACTTACAGCT 2156
DB |||||
1800 CCGTGAACCTGATCAGAGAGTACTTAACAGCCCTCTCTCAAGCTCACTTACAGCT 1859
QY |||||
2157 -TCTACTTAGTCCAGACGAGTCTGGAGACTCTGGCGGAGCCTACCAAGAACACTG 2215
DB |||||
1860 CTCTACTTAGTCCAGACGAGTCTGGAGACTCTGGCGGAGCCTACCAAGAACACTG 1919
QY |||||
2216 GACCGACCGGTGGTACCTCACCTTTACCGAGTCCGCGACACAGTGTGGGTCCGCGAC 2275
DB |||||
1920 GACCGACCGGTGGTACCTCACCTTTACCGAGTCCGCGACACAGTGTGGGTCCGCGAC 1979
QY |||||
2276 CAGACTAAGAACCTAGAACCTCGCTGGAAAGACCTTACAGTCTCTGCTGACACCCCT 2335
DB |||||
1980 CAGACTAAGAACCTAGAACCTCGCTGGAAAGACCTTACAGTCTCTGCTGACACCCCT 2039
QY |||||
2336 ACCGCCCTCAAGTAGAGCGGATCGCAGCTTGGATACACGCGCGCCACCTGAGGCTGCC 2395
DB |||||
2040 ACCGCCCTCAAGTAGAGCGGATCGC - GCTTGGATACACGCGCGCCACCTGAGAGCTGCC 2098
QY |||||
2396 GACCCCGGGGTGGACCACTCTCTAGACTGGCATGGGATGGAGTGTATCATCTCTTCT 2455
DB |||||
2099 GACCCCGGGGTGGACCACTCTCTAGACTGCC-----GGATCTCGAGG 2142
QY |||||
2456 TGGTAGAACAGCTACAGGTGTCACTCCGACATCCAGTACCCAGAGCCCAAGAGCC 2515
DB |||||
2143 ATCCACCACTAGGACCCCACTTAAATTTGGAATTCCTGACGCGCGGGGATCCACTAGT 2202
QY |||||
2516 TGAGCGCAGCGTGGGTGACAGAGTGACCATCACTGTAAAGCCAGTCAAGATGTGGGTA 2575
DB |||||
2203 CTAGAGCGAAATTAATTCGGTTATTTTCCACA-----TAA 2237

QY 2576 CTTCTGTAGCTTGGTACAGCAGAAAGCCAGGTAAAGCTTCAAAGCTGTCTATCTGA 2635
DB |||||
2238 TTGCGCTCTTTTGGCAATGTGAGGCGCCGGAACCTTGGCCCTGTCTTCTTGAAGCAAT 2297
QY |||||
2636 CATCACCCGCGACACTGGTGTGCAACAGATTCAGCGGTAGCGGTAGCGTACCGACT 2695
DB |||||
2298 CTTAGGGGTCTTTCCCTCTCGCCAAAGAAATGCAAGGTCTGTTGAATGTCTGTGAAGNA 2357
QY |||||
2696 TCACCTTCACCATCAGCAGCTCCAGCCAGAGGACATCGCCACCTTACTACTGCCACAT 2755
DB |||||
2358 GCAGTTCTCTGGAAGCTTCTTGAAGACAAACAGCTCTGTAGCGACCTT----- 2408
QY |||||
2756 ATAGCCTCTATCGGTCTGTTGCGCCAGAGGACCAAGGTGGAATCAAACAGAGTGGTCA 2815
DB |||||
2409 -----TGCAGGCGAGCGAAACCCCACTTGGCGACAGGTGCTCTGCG 2451
QY |||||
2816 GATCGGTGGATCGGCTCTGCTGCTCAGATCGGAGGTCCAACTGGTGGAGAGCGGTG 2875
DB |||||
2452 GCCAAAAGCCACGTGTATAGATACACCTGCAAGGCGGCAACCCAGTGGCCACGTG 2511
QY |||||
2876 GAGGTGTTGTCAACCTGCGCGCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 2935
DB |||||
2512 TGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCTATTTCAACAGGGGC 2571
QY |||||
2936 TCACACATATTTGATGATGTTGGGTGAGACAGGACCTTGAAGAGTCTTGAAGTGAATG 2995
DB |||||
2572 TGAAGGATGCGCAGAGGTACCCATTTGATGAGGATCTGATCTGGGCTCTGCTGACAT 2631
QY |||||
2996 GAGAAATTCATCCAGATAGTAGTACGATTAATGCTGCGCTCTCTAAAGGATAGATT 3055
DB |||||
2632 GCTTTACATGTTTGTAGTCGAGGTTAAAAAACGCTTAGGCCCCCGCAACACCGGGAG 2691
QY |||||
3056 CAATATCGCAGACAAACCGCAAGAACACATTTGTTCTCTGCAAAATGGACAGCCTGACAC 3115
DB |||||
2692 GGTTCCTTTGAAAACACCATGATATATG-----GGGATCCACCGG 2736
QY |||||
3116 AAGACACGGGTCTATTTTGTGCAAGCTTTTACTTTCGGCTTCCCTTGGTTTGTATT 3175
DB |||||
2737 TCGCACCATGTTGAGCAAGGCGGAGGAGTGTTCACCGGGTGG--TGCCCATCTGCT 2794
QY |||||
3176 GGGCCCAAGGACCCCGCTCAGCTCTCCAGTGTAGCCCAACCAAGCAGCGCGCGC 3235
DB |||||
2795 CGAGCTGAGCGGCGACGTAAACCGGCCACAGTTTCAAGCTGTCCGCGAGGCGGAGGCG 2854
QY |||||
3236 GACCAACAAACCGCGCCCACTCGCTCGCAGCCCTCTGCTGCGCCAGAGGCGG 3295
DB |||||
2855 TGCCACCTACGGCA-AGCTGACCTTGAAGTTCACTGACACCGCGGAGTGGCCGTGC 2913
QY |||||
3296 CTGCGCAGCGCGCGGGGCGCAGTGCACAGAGGGGCTGAGCTTCCCTTGGATCCCA 3355
DB |||||
2914 CTTGGCCACCTCGTGACACCTGACCTACCGCGTGCAGTGTCTCAGCGCTACCCCG 2973
QY |||||
3356 AACTCTGCTACTCTGTGATGGAATCTCTTCTATCTATGTTGTCTTCTACTGCTTGT 3415
DB |||||
2974 ACCACATGAAGCAGCAGCA-----CTTCTTCAAGTCCGCTATGCCGAAGGCTACG 3024
QY |||||
3416 TCCTGAGAGTGAAGTTCAAGCAGGAGCGCAGAGCCCGCGTACAGCAGCGCCAGAAC 3475
DB |||||
3025 TCCAGAGCGCACCATCTTCTTCAAGGACGCGCACTACAGACCCCGCGCGAGGTGA 3084
QY |||||
3476 AGCTCTATAACGAGCTCAATCTAGGACGAAGAGGAGTACGATGTTTGGACAGAGAC 3535
DB |||||
3085 AGTTTCGAGGGCGACACCTGTGTGAACCGCATCGAGCTGAAGGCGATCGACTTCAAGGAGG 3144
QY |||||
3536 GTGCGCGGACCTGAGATGGGGGAAAGCCGAGAGGAAGAAACCTCAGG---AAGGCC 3592
DB |||||
3145 ACGGCAACATCTCTGGGCAAGCTGGAGTCAAACTACAAAGCGCACAACTCTATATCA 3204
QY |||||
3593 TGTAACATGAACTCAGAAAGATTAAGTGGCGGAGGCTTACAGTGAGATTGGGATGAAG 3652
DB |||||
3205 TGGCCGACAGCAAGAAACCGCATCAAGGTGAATCTTCAAGATCCGCGCACAACTCGAGG 3264
QY |||||
3653 GCGAGCGCGGAGGGGCAAGGGGCGACGATGCGCTTTTACAGGGTCTCAGTACAGCCACA 3712

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2005, 08:04:19 ; Search time 14653 Seconds
(without alignments)
19882.885 Million cell updates/sec

Title: US-10-006-771B-1
Perfect score: 7654
Sequence: 1 agcttgcatgctgcaggt.....gtaaaacgacgcagtgcc 7654

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1055.8	13.8	1070	1 AJ281552	4A3A-P6F1
2	1018.2	13.3	1048	7 CO552396	ACLY4_50
3	974.4	12.7	1013	4 BM438846	IpLvr0015
4	947.8	12.4	1067	1 AU081137	AU081137 Mus.muscu
5	945.8	12.4	1265	9 AG435223	AG435223 Mus.muscu
6	944.4	12.3	1027	9 CL021190	CL021190 CH216-8A1
7	924	12.1	1004	1 AJ281480	AJ281480 4A3A-P4G8
8	916.4	12.0	1089	1 AU081124	AU081124 AU081124
9	909	11.9	1338	9 AG429657	AG429657 Mus.muscu
10	893.8	11.7	1049	9 CL021189	CL021189 CH216-8A1
11	889.8	11.6	928	7 CO487414	CO487414 GQ0227.B7
12	882.6	11.5	889	9 CL076017	CL076017 CH216-138
13	875.6	11.4	1028	9 CL021194	CL021194 CH216-8A1
14	871.4	11.4	925	6 CB686151	CB686151 Bn01b_020
15	865	11.3	1073	7 CF269652	CF269652 Fcylcol8
16	864.4	11.3	902	7 CR753463	CR753463 DKFZp469G
17	863.8	11.3	935	4 BG838279	BG838279 Gc01_10e0
18	857.6	11.2	885	9 CL076016	CL076016 CH216-138
19	831.8	10.9	1269	9 AG393068	AG393068 Mus.muscu
20	829	10.8	841	1 AL042026	AL042026 DKFZp434E
21	820.2	10.7	865	7 CK125894	CK125894 BES182411
22	819.6	10.7	1249	8 BZ572284	BZ572284 msh2_2572
23	815.4	10.7	1025	9 CL021193	CL021193 CH216-8A1
24	815	10.6	1169	9 AG332951	AG332951 Mus.muscu

C 25	813	10.6	1284	9	AG429476	Mus.muscu
C 26	812.6	10.6	854	4	BM438950	IpLvr0049
C 27	807.6	10.6	966	8	BZ570738	msh2_1513
C 28	806.6	10.5	1321	9	AG429698	Mus.muscu
C 29	802.4	10.5	1126	8	BZ577702	msh2_5533
C 30	797.4	10.4	856	7	CN823189	Oa.sp1bn
C 31	796	10.4	819	6	CD649375	CvGnd0008
C 32	785.6	10.3	1011	8	BZ576726	msh2_5071
C 33	785.4	10.3	1574	8	BZ572566	msh2_2693
C 34	783.6	10.2	820	7	CR753457	DKFZp469P
C 35	781	10.2	827	7	CN823902	Oa.sp1bn
C 36	780.8	10.2	806	4	BJ684280	BJ684280
C 37	779.8	10.2	846	7	CV468077	est_l_van
C 38	779.4	10.2	1163	1	AU081044	AU081044
C 39	778.6	10.2	833	6	CB686421	Bn01b_04f
C 40	777.8	10.2	800	1	AJ281449	4A3A-P4D5
C 41	772.4	10.1	954	1	AL044364	DKFZp434C
C 42	771.8	10.1	1245	9	AG448809	Mus.muscu
C 43	767.6	10.0	863	7	CF752100	TGDR9_Hum
C 44	762.8	10.0	1336	8	BZ575810	msh2_4637
C 45	762.4	10.0	789	6	CD280920	G44224.42

ALIGNMENTS

RESULT 1
AJ281552 1070 bp mRNA linear EST 30-JUN-2000
LOCUS 4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION AJ281552
VERSION AJ281552.1 GI:6929432
KEYWORDS EST.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
REFERENCE 1 (bases 1 to 1070)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
PUBMED 10841561
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .1070
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

```
Query Match      13.8%; Score 1055.8; DB 1; Length 1070;
Best Local Similarity 99.7%; Pred. No. 3.9e-290;
Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 5427 TTGCTGGCGTCTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAAAATCGACGCTCA 5486
Db 1 TTGCTGGCGTCTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAAAATCGACGCTCA 60

QY 5487 AGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCAAGCGTCTTCCCGCTGGAGC 5546
Db 61 AGTCAGAGGTGGCGAAACCCGACAGACTATAAGATACCAAGCGTCTTCCCGCTGGAGC 120

QY 5547 TCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTC 5606
Db 121 TCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTC 180

QY 5607 CTTCCGGAAGCGTGGCGCTTTCTCATAGCTCAAGCTAGCTAGTATCTCAGTTCGGTGTAG 5666
Db 181 CTTCCGGAAGCGTGGCGCTTTCTCATAGCTCAAGCTAGCTAGTATCTCAGTTCGGTGTAG 240

QY 5667 GTCGTTCCGTCACAGCTGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGGCC 5726
Db 241 GTCGTTCCGTCACAGCTGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGGCC 300

QY 5727 TTATCCGCTAACTATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCACTGGCA 5786
Db 301 TTATCCGCTAACTATCGTCTTGAGTCCAAACCGGTAAAGACAGACTTATCGCACTGGCA 360

QY 5787 CGAGCCACTGTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTG 5846
Db 361 CGAGCCACTGTGTAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTTCTTG 420

QY 5847 AAGTGGTGGCTTAACCTAGCGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 5906
Db 421 AAGTGGTGGCTTAACCTAGCGCTACACTAGAGGACAGTATTTGGTATCTGGCTCTGCTG 480

QY 5907 AAGCAGATTACTCTCGGAAAAAGAGTGTGTAGTCTTGTATCGGCAAAACACACCGCT 5966
Db 481 AAGCAGATTACTCTCGGAAAAAGAGTGTGTAGTCTTGTATCGGCAAAACACACCGCT 540

QY 5967 GGTAGCGGTGTTTTTTTTTTTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAA 6026
Db 541 GGTAGCGGTGTTTTTTTTTTTGTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAA 600

QY 6027 GAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACCTCAGTTAA 6086
Db 601 GAAGATCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAACCTCAGTTAA 660

QY 6087 GGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTTCAGCTAGATCTTTTAAATTTAAAA 6146
Db 661 GGGATTTTGGTTCATGAGATTATCAAAAAGGATCTTTCAGCTAGATCTTTTAAATTTAAAA 720

QY 6147 TGAAGTTTAAATCAATCAATGATATATAGTAAACTTGGTCTGACAGATTACCAATGC 6206
Db 721 TGAAGTTTAAATCAATCAATGATATATAGTAAACTTGGTCTGACAGATTACCAATGC 780

QY 6207 TTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGGTTCTATCATAGTTGCTGTA 6266
Db 781 TTAATCAGTAGGACCTATCTCAGCGATCTGTCTATTTGGTTCTATCATAGTTGCTGTA 840

QY 6267 CTCCTCGTGTGATAGTAACATACGATACGAGGAGGCTTACCATCTCGGCCCGCAGTGTGCA 6326
Db 841 CTCCTCGTGTGATAGTAACATACGATACGAGGAGGCTTACCATCTCGGCCCGCAGTGTGCA 900

QY 6327 ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATTAACCCAGCAGCC 6386
Db 901 ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTTATCAGCAATTAACCCAGCAGCC 960

QY 6387 GGAAGGCGGACGAGAGTGGTCTCTGCAACTTTATCCGCTTCATCCAGTCTATTAAAT 6446
Db 961 GGAAGGCGGACGAGAGTGGTCTCTGCAACTTTATCCGCTTCATCCAGTCTATTAAAT 1019

QY 6447 TGTTCGCGGAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTATTTTGGCGCAAC 6497
```

```
Db 1020 TGTTCGCGGAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTATTTGGCGCAAC 1070

RESULT 2
COS52396
LOCUS
DEFINITION
  Acly4_50 Sea lamprey Acly Petromyzon marinus cDNA, mRNA sequence.
ACCESSION
  COS52396
VERSION
  COS52396.1 GI:51800732
KEYWORDS
  EST.
SOURCE
  Petromyzon marinus (sea lamprey)
ORGANISM
  Petromyzon marinus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
  Petromyzontiformes; Petromyzontidae; Petromyzon.
REFERENCE
  1 (Bases 1 to 1048)
  Pancer, Z., Mayer, W.E., Klein, J. and Cooper, M.D.
  Prototypic T-cell receptor and CD4-like coreceptor expressed in
  lymphocytes of the agnathan sea lamprey
  Proc. Natl. Acad. Sci. U.S.A. 101, 13273-13278 (2004)
JOURNAL
  Contact: Pancer, Zeev
COMMENT
  Division of Developmental and Clinical Immunology
  The University of Alabama at Birmingham
  378 Wallace Tumor Institute, 1530 Third Avenue, South, Birmingham,
  AL 35294-3300
  Tel: 205-975-5812
  Fax: 205-975-7218
  Email: zpancer@uab.edu.
FEATURES
  Location/Qualifiers
  1..1048
    /organism="Petromyzon marinus"
    /mol_type="mRNA"
    /db_xref="taxon:7757"
    /cell_type="lymphocyte"
    /dev_stage="immune stimulated larvae"
    /clone_lib="Sea lamprey Acly"
    /note="Vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR
    subtracted cDNA libraries of immune stimulated larvae. All
    are single pass 5' or 3' sequences randomly cloned in
    pGEM-T Easy (Promega)."
```

```
ORIGIN
Query Match      13.3%; Score 1018.2; DB 7; Length 1048;
Best Local Similarity 99.2%; Pred. No. 2.3e-279;
Matches 1043; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

QY 5363 TAACGCGAGAAAGAACATGTGAGCAAAAGGCGCAGAAAAGCCAGGAAACCGTAAAAAGGC 5422
Db 1 TAACGAGGAAAGAACATGTGAGCAAAAGGCGCAGAAAAGCCAGGAAACCGTAAAAAGGC 60

QY 5423 CGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTCCCTGACGAGCATCAAAAAATCGACG 5482
Db 61 CGCGTTGCTGGCGTTTTTCCATAGGCTCCG-CCCGCTGACGAGCATCAAAAAATCGACG 119

QY 5483 CTCAGTTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGG 5542
Db 120 CTCAGTTCAGAGTGGCGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCTGG 179

QY 5543 AAGCTCCCTCGTGGCGCTCTCTCTGTTCCGACCTCCCGCTTACCGGATACCTGTCCGCTT 5602
Db 180 AAGCTCCCTCGTGGCGCTCTCTCTGTTCCGACCTCCCGCTTACCGGATACCTGTCCGCTT 239

QY 5603 TCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGT 5662
Db 240 TCTCCCTTCGGGAAGSGTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGT 299

QY 5663 GTAGTGTTCGCTCCAGCTGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTG 5722
Db 300 GTAGTGTTCGCTCCAGCTGGGCTGTGTGACAGAACCCCGCTTACGCCGACCGCTG 359

QY 5723 CGCCTTATCCGGTAACTATCGTCTTTGAGTCCAAACCCCGGTAAAGACGAGTATTCGCCACT 5782
Db 360 CGCCTTATCCGGTAACTATCGTCTTTGAGTCCAAACCCCGGTAAAGACGAGTATTCGCCACT 419
```

QY	5783	GGCAGCAGCCACTGGTAAACAGCAGATTACGACAGACGAGGTATGTAGCGCGTCTACAGAGTT	5842
Db	420	GGCAGCAGCCACTGGTAAACAGCAGATTACGACAGACGAGGTATGTAGCGCGTCTACAGAGTT	479
QY	5843	CTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCGTCT	5902
Db	480	CTTGAAGTGGTGGCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTCGCGTCT	539
QY	5903	GCTGAAGCCAGTTACCTTTCCGAAAAAGAGTTGGTAGTCTTTGTATCCGCGCAAAACAAACAC	5962
Db	540	GCTGAAGCCAGTTACCTTTCCGAAAAAGAGTTGGTAGTCTTTGTATCCGCGCAAAACAAACAC	599
QY	5963	CGCTGGTAGCGGTGGTTTTTTTGGTTTGAACAGCAGATTACGGCGCAGAAAAAAGGATC	6022
Db	600	CGCTGGTAGCGG--TGTTTTTTTGGTTTGAACAGCAGATTACGGCGCAGAAAAAAGGATC	657
QY	6023	TCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGC	6082
Db	658	TCAAGAGATCCTTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGC	717
QY	6083	TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTA	6142
Db	718	TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTA	777
QY	6143	AAAATGAAGTTTAAATCAATCTAAAGTATATATAGATTAACCTTGGTCTGACAGTTACCA	6202
Db	778	AAAATGAAGTTTAAATCAATCTAAAGTATATATAGATTAACCTTGGTCTGACAGTTACCA	837
QY	6203	ATGCTTAATCAGTCAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGC	6262
Db	838	ATGCTTAATCAGTCAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGC	897
QY	6263	CTGACTCCCCGTCGTGTAGATAAATACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC	6322
Db	898	CTGACTCCCCGTCGTGTAGATAAATACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC	957
QY	6323	TGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACACGCC	6382
Db	958	TGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGATTATTCAGCAATAAACACGCC	1017
QY	6383	AGCCGGAGGGCCGAGCGCAGAGTGGTCT	6413
Db	1018	AGCCGGAGGGCCGAGCGCAGAGTGGTCT	1048
RESULT 3			
LOCUS	BM438846/c		
DEFINITION	IP19v00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA	1013 bp	linear EST 31-JAN-2002
ACCESSION	BM438846		
VERSION	BM438846.1		
KEYWORDS	EST.		
SOURCE	Ictalurus punctatus (channel catfish)		
ORGANISM	Ictalurus punctatus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes; Ictaluridae; Ictalurus.		
AUTHORS	Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.		
TITLE	Transcriptome of channel catfish (Ictalurus punctatus): initial analysis of expressed sequence tags from the liver		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Liu ZJ		
	The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences		
	Auburn University		
	203 Swingle Hall, Auburn University, Auburn, AL 36849, USA		
	Tel: 334 844 4054		
	Fax: 334 844 9208		
	Email: zliu@acesag.auburn.edu		

Db 174 CAGAACTTTAAAGTGCTCATCATTTGGAAAGAGTTCTTCGGGGCGAAAACTCTCAAGGAT 115
 QY 6898 CTTACCGCTGTTGAGATCCAGTTCAGTGAACCCACTCGTGCACCACTGATCTTCAGC 6957
 Db 114 CTTACCGCTGTTGAGATCCAGTTCAGTGAACCCACTCGTGCACCACTGATCTTCAGC 55
 QY 6958 ATCTTTTACTTTTCCACGAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGC 7011
 Db 54 ATCTTTTACTTTTCCACGAGCTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGC 1

RESULT 4
 AU081137
 LOCUS
 DEFINITION
 AU081137 Oncorhynchus mykiss Kidney infected by infectious
 hematopoietic necrosis virus
 mRNA sequence.
 AU081137.1 GI:6431485

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oncorhynchus mykiss (rainbow trout)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 1 (bases 1 to 1067)
 Kono, T., Sakai, M. and Iapetra, S.B.
 Expressed Sequence tag Analysis of Kidney and Gill Tissues from
 Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
 Hematopoietic Necrosis Virus
 Mar. Biotechnol. 2 (5), 493-498 (2001)
 Contact: Masahiro Sakai
 Faculty of Agriculture
 Miyazaki University
 1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
 Email: m.sakai@cc.miyazaki-u.ac.jp.
 Location/Qualifiers

FEATURES
 source
 1..1067
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="K12"
 /tissue_type="Kidney"
 /clone_lib="Oncorhynchus mykiss Kidney infected by
 infectious hematopoietic necrosis virus"
 /note="common name:rainbow trout ; infected by infectious
 hematopoietic necrosis virus"

ORIGIN
 Query Match 12.4%; Score 947.8; DB 1; Length 1067;
 Best Local Similarity 99.1%; Pred. No. 3.5e-259;
 Matches 974; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 5026 COTATCATGTCATAGCTGTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 5085
 Db 87 CGTATCATGTCATAGCTGTTCTCTGTGTAATTTGTTATCCGCTCACAATTCACACA 146
 QY 5086 ACATACGAGCCGGAAGCATAAAGTGAAGCCCTGGGGTGCCCTAATGAGTGAGCTAACTCA 5145
 Db 147 ACATACGAGCCGGAAGCATAAAGTGAAGCCCTGGGGTGCCCTAATGAGTGAGCTAACTCA 206
 QY 5146 CATTAAATTTGGCTTGGCTCACTGCCCGCTTCCAGTCGGGAACCTGTGCGCAGCTGC 5205
 Db 207 CATTAAATTTGGCTTGGCTCACTGCCCGCTTCCAGTCGGGAACCTGTGCGCAGCTGC 266
 QY 5206 ATTAATGAATCGCCCAACCGCGGGAGAGCGGTTTTCGGTATTGGCGCTCTTCGCGTT 5265
 Db 267 ATTAATGAATCGCCCAACCGCGGGAGAGCGGTTTTCGGTATTGGCGCTCTTCGCGTT 326
 QY 5266 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 5325
 Db 327 CCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 386

QY 5326 CAAGGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGACGAGGAAGAACATGTGAG 5385
 Db 387 CAAGGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGACGAGGAAGAACATGTGAG 446
 QY 5386 CAAGGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGACGAGGAAGAACATGTGAG 5445
 Db 447 CAAGGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGACGAGGAAGAACATGTGAG 506
 QY 5446 GGCTCCGCGCCCTCGACGAGCATCACAAAATCAGCGCTCAAGTCAGAGTGGCGGAAC 5505
 Db 507 GGCTCCGCGCCCTCGACGAGCATCACAAAATCAGCGCTCAAGTCAGAGTGGCGGAAC 566
 QY 5506 CGACGAGTACTATAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTCCGCTCTCTG 5565
 Db 567 CGACGAGTACTATAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTCCGCTCTCTG 626
 QY 5566 TTCCGACCTCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGC 5625
 Db 627 TTCCGACCTCCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGC 686
 QY 5626 TTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCGCTCCAGCTGG 5685
 Db 687 TTCTCATAGCTACGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCGCTCCAGCTGG 746
 QY 5686 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGCTTAACCTATCGTC 5745
 Db 747 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGCTTAACCTATCGTC 806
 QY 5746 TTGAGTCCAAACCCCGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 5805
 Db 807 TTGAGTCCAAACCCCGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 866
 QY 5806 TTAGCAGGCGAGGTATGTAGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGC 5865
 Db 867 TTAGCAGGCGAGGTATGTAGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGC 926
 QY 5866 GCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAA 5925
 Db 927 GCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAA 986
 QY 5926 AAAGAGTTGGTGTAGCTTTGATCCGGCAAAACCAACCGCTGGTGGTGGTGGTGGTGGT 5985
 Db 987 AAAGAGTTGGTGTAGCTTTGATCCGGCAAAACCAACCGCTGGTGGTGGTGGTGGTGGT 1045
 QY 5986 TTTCGAGCAGCAGATTAACGGC 6008
 Db 1046 TTTCG-AGCAGCAGATTAACGGC 1067

RESULT 5
 AG435223/c
 LOCUS
 DEFINITION
 AG435223 1265 bp DNA linear GSS 03-JUN-2004
 Mus musculus molossinus DNA, clone:MSMg01-313D17.TJ, genomic survey
 sequence.
 ACCESSION
 AG435223
 AG435223.1 GI:48078286
 KEYWORDS
 GSS.
 SOURCE
 Mus musculus molossinus
 Mus musculus molossinus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 BAC end Sequences of Library MSMg01
 Unpublished
 2 (bases 1 to 1265)
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
 Direct Submission
 Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@sc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)

6706 TACTGTATGTCATCGTAAGATGCTTTTCTGACTGCTGAGTACTCAACCAAGTCATT 6765
 965 TACTGTATGTCATCGTAAGATGCTTTTCTGACTGCTGAGTACTCAACCAAGTCATT 906
 6766 CTGAGATAGTGTATGCGGCGACCGAGTGTCTTTGCGCGGCGTCAATACCGGATAATAC 6825
 905 CTGAGATAGTGTATGCGGCGACCGAGTGTCTTTGCGCGGCGTCAATACCGGATAATAC 846
 6826 CCGGCCATACGACAACTTTAAAGTGTCTATCATTTGAAACGTTTCTTCGGGCGGAAA 6885
 845 CCGGCCATACGACAACTTTAAAGTGTCTATCATTTGAAACGTTTCTTCGGGCGGAAA 786
 6886 ACTCTCAAGGATCTTACCGCTGTTCGAGATCCAGTGTGATGTAACCCACTCGTGCAACCAA 6945
 785 ACTCTCAAGGATCTTACCGCTGTTCGAGATCCAGTGTGATGTAACCCACTCGTGCAACCAA 726
 6946 CTGATCTTTCAGCATCTTTTACTTTTACCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGCA 7005
 725 CTGATCTTTCAGCATCTTTTACTTTTACCAGCGCTTTCTGGGTGAGCAAAAACAGGAAGCA 666
 7006 AAATCCGCAAAAAGGGAATTAAGGCGACGCGAAATGTTGAATACATCTCTTCT 7065
 665 AAATCCGCAAAAAGGGAATTAAGGCGCGACACGGAATGTTGAATACATCTCTTCT 606
 7066 TTTTCAATATATTGAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGA 7125
 605 TTTTCAATATATTGAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGA 546
 7126 ATGTATTTAGAAAAATAAACAATAGGGTTCGCGGCACTTTCCCGGAAAGTCCAC 7185
 545 ATGTATTTAGAAAAATAAACAATAGGGTTCGCGGCACTTTCCCGGAAAGTCCAC 486
 7186 TGACGTCTAAGAAACATTTATCATGACATTAACCTATAAATAAGCGGTATCACGAG 7245
 485 TGACGTCTAAGAAACATTTATCATGACATTAACCTATAAATAAGCGGTATCACGAG 426
 7246 GCCCTTTCTGCTCGCGGCTTCGCGTATGACGGTGAACCTCTGACACATGCACTCCC 7305
 425 GCCCTTTCTGCTCGCGGCTTCGCGTATGACGGTGAACCTCTGACACATGCACTCCC 366
 7306 GGAGACGCTCACAGCTTGTCTGTAAGCGGATGCGGCGACGACAAAGCCGTCAGGGCG 7365
 365 GGAGACGCTCACAGCTTGTCTGTAAGCGGATGCGGCGACGACAAAGCCGTCAGGGCG 306
 7366 GTACGCGGCTGTTCGCGGCTGTTCGCGTATGACGGTGAACCTCTGACACATGCACTCCC 7425
 305 GTACGCGGCTGTTCGCGGCTGTTCGCGTATGACGGTGAACCTCTGACACATGCACTCCC 246
 7426 ACTGAGGTGACCATATGCGGTGTGAATACCGCACAGATCGGTAAAGGAAATACCG 7485
 245 ACTGAGGTGACCATATGCGGTGTGAATACCGCACAGATCGGTAAAGGAAATACCG 186
 7486 CATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGC 7545
 185 CATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGCGCATCGGTGCGGC 126
 7546 CTCTTCGCTATTAACGCGAGTGGCGAAAGGGGATGTGCTGCAAGCGCATTAAGTTGGGT 7605
 125 CTCTTCGCTATTAACGCGAGTGGCGAAA - GGCGATGTGCTGCAAGGCGCATTAAGTTGGG- 68
 7606 AACGCCGGGTTTCCCGAGTCAGACGTTGTAAA 7639
 67 AACGCCGGGTTTCCCGAGTCAGACGTTGTAAA 34

RESULT 7
 AJ281480
 LOCUS
 DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
 gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
 AJ281480
 ACCESSION
 VERSION AJ281480.1 GI:6929360

EST.
 Anopheles gambiae (African malaria mosquito)
 Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 Anopheles.
 1 (bases 1 to 1004)
 Dinopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
 Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.
 and Kafatos, F.C.
 Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
 20300950
 10841561
 Contact: Dimopoulos G
 Fotis C. Kafatos laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 Location/Qualifiers
 1..1004
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="4A r/r"
 /db_xref="taxon:7165"
 /clone="4A3A-P4G8"
 /lab_host="E. coli DH10B"
 /clone_lib="Anopheles gambiae immune competent 4A3A"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
 (1996) : Normalization and Subtraction: Two approaches To
 Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match 12.1%; Score 924; DB 1; Length 1004;
 Best Local Similarity 99.0%; Pred. No. 2.3e-252;
 Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;
 QY 5398 AAAAGGCCAGAAACCGTAAAAAGGCCGGTGTCTGCGGCTTTTCCATAGGCTCCGCCCC 5457
 DB 1 AAAAGGCCAGAAACCGTAAAAAGGCCGGTGTCTGCGGCTTTTCCATAGGCTCCGCCCC 60
 QY 5458 CTGACGAGCATCACAAAATCGAGCTCAAGTCAAGGTGCGAAACCCGACAGGACTAT 5517
 DB 61 CTGACGAGCATCACAAAATCGAGCTCAAGTCAAGGTGCGAAACCCGACAGGACTAT 120
 QY 5518 AAAGATACAGGCGTTCCTCCCTCGAAGCTCCCTGCTGCGCTCTCTCTTCGACCCCTGC 5577
 DB 121 AAAGATACAGGCGTTCCTCCCTCGAAGCTCCCTGCTGCGCTCTCTCTTCGACCCCTGC 180
 QY 5578 CGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTAGCT 5637
 DB 181 CGCTTACCGGATACCTGTCGCCCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCTAGCT 240
 QY 5638 CACGCTGTAGGTATCTCAGTTTCGGGTGAGTTCGCTCCAAAGCTGGGTGTGTGACG 5697
 DB 241 CACGCTGTAGGTATCTCAGTTTCGGGTGAGTTCGCTCCAAAGCTGGGTGTGTGACG 300
 QY 5698 AACCCCGGTTTCAGCCCGACCGCTCGGCTTTATCGGTAACTATCGTCTTGAAGTCAAACC 5757
 DB 301 AACCCCGGTTTCAGCCCGACCGCTCGGCTTTATCGGTAACTATCGTCTTGAAGTCAAACC 360
 QY 5758 CGGTAAAGACAGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTTAGCAGAGCA 5817
 DB 361 CGGTAAAGACAGACTTATCGCACTGGCAGCAGCCACTGGTAACAGGATTTAGCAGAGCA 420
 QY 5818 GGTATGTAGGCGGTGTACAGAGTCTTGAAGTGTGTGCGCTTAACCTACGGCTACACTAGNA 5877

```

|||||
421 GGTATGAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGCGTACACTAGAA 480
Qy
5878 GGCAGTATTGGTATCTGCGCTCTGCTGAAGCGAGTTACCTTCGGAAAAAGAGTTGGTA 5937
Db
481 GAACAGTATTGGTATCTGCGCTCTGCTGAAGCGAGTTACCTTCGGAAAAAGAGTTGGTA 540
Qy
5938 GCTCTTGATCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGGTTGCAAGCAGC 5997
Db
541 GCTCTTGATCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGGTTGCAAGCAGC 600
Qy
5998 AGATTACGCGCAGAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG 6057
Db
601 AGATTACGCGCAGAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG 660
Qy
6058 ACCTCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGA 6117
Db
661 ACCTCTCAGTGGAAACGAAACTCAGTTTAAAGGATTTTGGTCATGAGATTAATCAAAAAGGA 720
Qy
6118 TCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATG 6177
Db
721 TCTTACCTAGATCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGTATATATG 780
Qy
6178 AGTAAACTCTGCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCT 6237
Db
781 AGTAAACTCTGCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTCAGCGATCT 839
Qy
6238 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGCTGCTGATAGATAAATCAATGATACGGG 6297
Db
840 GTCTATTTCGTTTCATCATAGTTGCTGACTCCCGCTGCTGATAGATAAATCAATGATACGGG 898
Qy
6298 AGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGAGACCCACGCTCACCGGCTC 6357
Db
899 AGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGAGGCGGAGCGGAGGCGGAGGCGGAGG 6408
Qy
6358 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 6408
Db
956 CAGATTATCAGCAATAAACCCAGCCAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 1004

RESULT 8
AU081124 1089 bp mRNA linear EST 30-JUL-2002
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG/12,
mRNA sequence.
ACCESSION AU081124.1 GI:6431472
VERSION 1
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Kono,T., Sakai,M. and LaPatra,S.E.
TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakukenibandai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
Location/Qualifiers
FEATURES
source
1.1089
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="KG/12"
/tissue_type="kidney"
/clone_lib="Oncorhynchus mykiss Kidney infected by

```

```

infectious hematopoietic necrosis virus"
/notes="common name:rainbow trout ; infected by infectious
hematopoietic necrosis virus"

Query Match 12.0%; Score 916.4; DB 1; Length 1089;
Best Local Similarity 97.3%; Pred. No. 3.5e-250;
Matches 975; Conservative 0; Mismatches 21; Indels 6; Gaps 4;

ORIGIN
Qy 5026 CGTAATCATGGTTCATAGCTGTTTCCCTGTGTGAAATTTGTTATCGGCTCACAAATTCACACA 5085
Db 88 CGTAATCATGGTTCATAGCTGTTTCCCTGTGTGAAATTTGTTATCGGCTCACAAATTCACACA 147
Qy 5086 ACATAGAGCGCGAAGCATAAAGCTGTAAAGCCTTGGGGTGCCTAATGAGTGAGCTAACTCA 5145
Db 148 ACATAGAGCGCGAAGCATAAAGCTGTAAAGCCTTGGGGTGCCTAATGAGTGAGCTAACTCA 207
Qy 5146 CATTAAATTCGTTGGCTCCTACCTCCCGCTTTCCAGTCGGGAAACCTGTCGTGCGAGCTGC 5205
Db 208 CATTAAATTCGTTGGCTCCTACCTCCCGCTTTCCAGTCGGGAAACCTGTCGTGCGAGCTGC 267
Qy 5206 ATTAATGAATTCGGCCAAACGCGCGGGGAGAGCGGTTTGGCGTATTGGGCGCTCTTCCGCTT 5265
Db 268 ATTAATGAATTCGGCCAAACGCGCGGGGAGAGCGGTTTGGCGTATTGGGCGCTCTTCCGCTT 327
Qy 5266 CTTGCTCTCACTGACTCGTGGCTCGGTTCGTTTCGCTGCGGCGAGCGGTATCAGCTCACT 5325
Db 328 CTTGCTCTCACTGACTCGTGGCTCGGTTCGTTTCGCTGCGGCGAGCGGTATCAGCTCACT 387
Qy 5326 CANAGCGGTATATACGGTTATCCACAGNATCAGGGGATACGCGAGGAAAGCAATGTGAG 5385
Db 388 CAAAGCGGTATATACGGTTATCCACAGNATCAGGGGATACGCGAGGAAAGCAATGTGAG 447
Qy 5386 CAAAAGCCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTTGTGGCGCTTTTCCATA 5445
Db 448 CAAAAGCCAGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTTGTGGCGCTTTTCCATA 507
Qy 5446 GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC 5505
Db 508 GGCTCCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAAGAGTGGCGAAACC 567
Qy 5506 CGACAGACTATATAAGATACCGAGCTTTCCCGCTGGAGCTCCCTGTCGGCTCTCTG 5565
Db 568 CGACAGACTATATAAGATACCGAGCTTTCCCGCTGGAGCTCCCTGTCGGCTCTCTG 627
Qy 5566 TTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGC 5625
Db 628 TTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGC 687
Qy 5626 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 5685
Db 688 TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 747
Qy 5686 GCTGTGTGACGAACCCCGCTTACCGCGACCGCTGCGCTTATCCGGTAACTATCGTC 5745
Db 748 GCTGTGTGACGAACCCCGCTTACCGCGACCGCTGCGCTTATCCGGTAACTATCGTC 807
Qy 5746 TTGAGTCCAAACCCCGGTAAGACAGCTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGA 5805
Db 808 TTGAGTCCAAACCCCGGTAAGACAGCTTATCGCACCTGGCAGCAGCCACTGGTAAACAGGA 867
Qy 5806 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAG-TGGTGGCCCTAACTAC 5864
Db 868 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTTTGTGGCTTAACACTAC 927
Qy 5865 GGCTACATAGAAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTGTACCTCGGA 5924
Db 928 GGCTACATAGAAAGGACAGTATTTTGTATCTGCGCTCTGCTGGAGCCAGTGTACCTCGGA 987
Qy 5925 ABAAGAGTTGGTAGCTCTT-GATCCGCGCAACAAACCA---CCGCTGGTAGCGGTGTTT 5980
Db 988 ABAAGAGTTGGTAGCTCTTGGATCCCGCAACAAACCAACCCGCTGGGAAGGGGGTGTCT 1047

```

```

Qy 5981 TTTTCTTTGCAAGCA-GCAGATTACCGCAGAAAAAAGGAT 6021
Db 1048 TTTTCTTTGCAAGCAGCAGATTACCGCGAAAAAACGGAT 1089

RESULT 9
AG429657/c
LOCUS
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-304F12.TJ, genomic survey
sequence.
ACCESSION AG429657
VERSION AG429657.1 GI:48072720
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1338)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI.
R.Site 2 : EcoRI.
FEATURES
Location/Qualifiers
1. 1338
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-304F12.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 11.9%; Score 909; DB 9; Length 1338;
Best Local Similarity 92.1%; Pred.No.5.1e-248;
Matches 1025; Conservative 0; Mismatches 77; Indels 11; Gaps 6;

Qy 6540 TTGGTATCGCTTCATTCAGCTCCGGTTCCTCCCAACGATCAAGGCGAGTTACATGATCCCCCA 6599
Db 1191 TTTGTTTGTAAAGTTCTCCTACTCGGTTCCCAAGGATCAAGGGGTTCCCTGTTCCTCCCTCC- 1133

Qy 6600 TGTGTGCAAAAACGGTTAGTCTCCTTCGGTCTCCGATCGTTGTCAGAAAGTAAAGTGG 6659
Db 1132 ---TTTGCAAAAAGGGGTTCTTCTTCGTCCTCCNATTC--TTTCCAAAAGTTAAATTGC 1078

Qy 6660 CCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCTCTTACTGTGCATGCCAT 6719
Db 1077 CGCA--GTTTTCTCCTCAGGTAGGCGAGCCCGCTTATTCTTCTACTGCCACCCAT 1020

Qy 6720 CCGTAAGATGC-TTTTCTGTGACTGGTGAGTACTCAA-CCAAGTCATTCTGAGAAATAGTG 6777
Db 1019 CCATAAGATGCTTTTCTGTGCGGGGGGAGTACTCAACCCCAAGTCTTCTCTGAGAAATAGTG 960

```

```

Qy 6778 TATCGCGCACCAGTTGCTCTTTCGCGGCGTCAATACGGGATAATACCGCGCCACATAG 6837
Db 959 TATCGCGCACCAGTTGCTCTTTCGCGGCGTCAATACGGGATAATACCGCGCCACATAG 900

Qy 6838 CAGAACTTTTAAAGTGTCTCATTCATTGGAAAAAGTTTTCGGGGCGAAAAAATCTCAAGAT 6897
Db 899 CAGAACTNTTAAAGTGTCTCATTCATGGGAAAAAGTTTTCGGGGCGAAAAAATCTCAAGAT 840

Qy 6998 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTGCTGTCACCACTGATCTTTCAGC 6957
Db 839 CTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTGCTGTCACCACTGATCTTTCAGC 780

Qy 6958 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATCGGCAAA 7017
Db 779 ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATCGGCAAA 720

Qy 7018 AAAGGGAATAGGGCGACACACGAAATGTTGAATACTCATACTCTTTCCTTTTCAATATTA 7077
Db 719 AAAGGGAATAGGGCGACACGAAATGTTGAATACTCATACTCTTTCCTTTTCAATATTA 660

Qy 7078 TTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGTAATGTATTAGAA 7137
Db 659 TTGAAGCATTTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGTAATGTATTAGAA 600

Qy 7138 AAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCCACCTGACGCTTAA 7197
Db 599 AAATAAACAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCCACCTGACGCTTAA 540

Qy 7198 AACCATTTATTCATGATCACTAACTTAACTATAAAATAGCGGTATCACAGAGCCCTTTCGTCT 7257
Db 539 AACCATTTATTCATGATCACTAACTTAACTATAAAATAGCGGTATCACAGAGCCCTTTCGTCT 480

Qy 7258 CGCGCGTTTCGGTGATGACGGTGAAAACTCTGACATGACGCTCCCGGAGACGGTCAC 7317
Db 479 CGCGCGTTTCGGTGATGACGGTGAAAACTCTGACATGACGCTCCCGGAGACGGTCAC 420

Qy 7318 AGCTTTGCTGTAAAGCGGATCCGGGAGCAGACAAAGCCGTCAGGCGCGTCAGCGGGTGT 7377
Db 419 AGCTTTGCTGTAAAGCGGATCCGGGAGCAGACAAAGCCGTCAGGCGCGTCAGCGGGTGT 360

Qy 7378 TGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCA 7437
Db 359 TGGCGGGTGTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCA 300

Qy 7438 CCATATCGGTGTGAAATACCGCAAGATGCTTAAGGAGAAAAATACCGCATCAGCGGCCA 7497
Db 299 CCATATCGGTGTGAAATACCGCAAGATGCTTAAGGAGAAAAATACCGCATCAGCGGCCA 240

Qy 7498 TTCCGCCATTACGGCTGCGCACTGTTGGGAGGGCGATCGGTGCGGGCTCTTCGCTATT 7557
Db 239 TTCCGCCATTACGGCTGCGCACTGTTGGGAGGGCGATCGGTGCGGGCTCTTCGCTATT 180

Qy 7558 ACGCAGCTGGCGAAAGGGGATGCTGCTGCAAGCGGATTAAAGTTGGGTAAACGCCA-GGGT 7616
Db 179 ACGCAGCTGGCGAAAGGGGATGCTGCTGCAAGCGGATTAAAGTTGGGTAAACGCCAGGGT 120

Qy 7617 TTTCCCATGACGACGTTGTAAACACGACGCCA 7649
Db 119 TTTCCCATGACGACGTTGTAAACACGACGCCA 87

RESULT 10
CL021189
LOCUS
DEFINITION CL021189 1049 bp DNA linear GSS 31-DEC-2003
CH216-8A14 RM1.1 CH216 xenopus tropicalis genomic clone CH216-8A14,
genomic survey sequence.
ACCESSION CL021189
VERSION CL021189.1 GI:40463002
KEYWORDS GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Qy	5626	TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGGTTTCGCTCCAAGCTGG	5685
Db	712	TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGGTTTCGCTCCAAGCTGG	771
Qy	5686	GCTGTGTGCAGCAACCCCGCTTCAGCCCGACCGCTCGCGCTTATCCGGTAACTATCGTC	5745
Db	772	GCTGTGTGCAGCAACCCCGCTTCAGCCCGACCGCTCGCGCTTATCCGGTAACTATCGTC	831
Qy	5746	TTGAGTCCAAACCCGGTAAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGA	5805
Db	832	TTGAGTCCAAACCCGGTAAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGA	891
Qy	5806	TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAAGTGGTGCCTAACTACG	5865
Db	892	TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTTGAAAGTGGTGCCTAACTACG	951
Qy	5866	GCTACACTAGAAGGAC-AGTATTGGTATCTCGCTCGCTGCTGAAGCAGCTTACG-TTCGG	5923
Db	952	GCTACACTAGAAGCAAGATTTGTTTATCTGGGCTCTGCTGAACAGTTACGCTTTCCG	1011
Qy	5924	AAAAAGATTGGTAGCTCTTTGATCGGCAACAAACC	5960
Db	1012	AAAAAAGTTGGTAGCTCTTTGATCGGCAACAAACC	1048
RESULT 11			
LOCUS	CO487414	928 bp	linear
DEFINITION	GQ0227.B7.1_K04_GQ022: ROOT XYLEM - mature trees Picea glauca cDNA clone GQ0227.B7_K04_3', mRNA sequence.		
ACCESSION	CO487414		
VERSION	CO487414.1	GI:50174316	
KEYWORDS	EST.		
SOURCE	Picea glauca (white spruce)		
ORGANISM	Picea glauca		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.		
AUTHORS	1 (bases 1 to 928) Morency M.-J., Cooke J., Pavy N., Parsons L., Paule C., Seguin A., Retzel E., Butterfield Y., Barber S., Yang G., Stott J., Siddiqui A., Holt R., Marra M. and Mackay J. Arborea EST sequencing in Picea glauca (white spruce) Unpublished (2004)		
TITLE	Contact: John Mackay		
JOURNAL	Centre de Recherche en Biologie Forestiere		
COMMENT	Universite Laval Pavillon Charles-Eugene Marchand, Quebec, Canada G1K 7P4 Fax: 418 656 7493 Email: jmackay@rsvs.ulaval.ca Center for Computational Genomics and Bioinformatics (CCGB), University of Minnesota, MN id Identifier: MN5286152 Clone ID: GQ0227.B7_K04 Clones available through: John Mackay, Ph. D. Professeur adjoint -Assistant professor EMAIL: jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center) Universite Laval Quebec, Quebec CANADA G1K 7P4 Plate: 7.B7 row: 04 column: K Seq primer: T7 Primer.		
FEATURES			
source	Location/Qualifiers		
	1..928		
	/organism="Picea glauca"		
	/mol_type="mRNA"		
	/strain="Three trees Unknown"		
	/db_xref="taxon:3330"		
	/clone="GQ0227.B7_K04"		
	/sex="Hexamphrodite"		
	/tissue_type="Differentiating xylem from roots 1 cm in diameter or larger"		
	/dev_stage="Non-lignified xylem tissues from larger roots in early part of growing season (June)."		
	/lab_host="E. coli DH10B cells"		
	/clone_lib="GQ022: ROOT XYLEM - mature trees"		

/note="Organ: Roots from 9 year old trees measuring approximately 4.5 m tall, and 10 cm in diameter.; Vector: pBluescript II SK (+) XR; Site 1: Eco-RI; Site 2: Xho-I; cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (in vitro) for propagation"

```
ORIGIN
Query Match      11.6%; Score 889.8; DB 7; Length 928;
Best Local Similarity 96.0%; Pred. No. 1.4e-242;
Matches 891; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 5438 TTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGACCTCAAGCTCAGAGTG 5497
Db 1 TTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGACCTCAAGTNNNNNG 60

Qy 5498 GCGAACCCGACGAGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTGTCG 5557
Db 61 GCGAACCCGACGAGACTATAAGATACAGGCGTTTCCCTCGAAGCTCCCTGTCG 120

Qy 5558 CTCTCCTCTCCGACCTCGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAG 5617
Db 121 CTCTCCTCTCCGACCTCGCGCTTACCGGATACCTGTCGCTTCTCCCTTCGGGAG 180

Qy 5618 CGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTTCGCTC 5677
Db 181 CGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTTCGGGTAGGTTCGCTC 240

Qy 5678 CAAGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCGGTAA 5737
Db 241 CAAGCTGGGCTGTGTGACGAAACCCCGTTTCAGCCGACCGCTGCGCTTATCGGTAA 300

Qy 5738 CTATCGTCTTGAGTCAACCCGGTAAGACACGACTTATCGCACTGGCAGCAGCACTGG 5797
Db 301 CTATCGTCTTGAGTCAACCCGGTAAGACACGACTTATCGCACTGGCAGCAGCACTGG 360

Qy 5798 TAAACAGGATTAGCAGAGCGAGTATGTAGGCGTGTCTACAGAGTTCCTTGAAGTGGTGCC 5857
Db 361 TAAACAGGATTAGCAGAGCGAGTATGTAGGCGTGTCTACAGAGTTCCTTGAAGTGGTGCC 420

Qy 5858 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGTCTGAAAGCCAGTTAC 5917
Db 421 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGCTCTGTCTGAAAGCCAGTTAC 480

Qy 5918 CTTCCGAAAAGAGTTGTAGTCTTTGATTCGGGCAACAAACCACCGCTGTAGCGGTGG 5977
Db 481 CTTCCGAAAAGAGTTGTAGTCTTTGATTCGGGCAACAAACCACCGCTGTAGCGGTGG 540

Qy 5978 TTTTTCCTGTTGCAAGCAGCAGATTACCGCGAGAAAAAAGATCTCAAGAAGATCCCTTT 6037
Db 541 TTTTTCCTGTTGCAAGCAGCAGATTACCGCGAGAAAAAAGATCTCAAGAAGATCCCTTT 600

Qy 6038 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGT 6097
Db 601 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTTGGT 660

Qy 6098 CATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTAAGATGAAGTTTAA 6157
Db 661 CATGAGATTATCAAAAAGGATCTTACCTAGATCTCTTTTAAATTAAGATGAAGTTTAA 720

Qy 6158 ATCAATCTAAAGTATATAGTAACTTGGTCTGACAGATTACCAATGCTTAAATCAGTGA 6217
Db 721 ATCAATCTAAAGTATATAGTAACTTGGTCTGACAGATTACCAATGCTTAAATCAGTGA 780

Qy 6218 GGCACCTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGT 6277
Db 781 GGCACCTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGT 840

Qy 6278 GTAGATAACTAGCATACGGGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGCG 6337
Db 841 GTNNNTNACTAGCATACGGGAGGCTTACCATCTGNNNNNNGTGTGCAATGATACCGCG 900
```

```
Qy 6338 AGACCCACGCTCACCGCTCCAGATTTA 6365
Db 901 AGACCCACGCTNNNNNNNNNNNNNNNTTNA 928

RESULT 12
LOCUS CL076017/c
DEFINITION CH216-138F20_Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-138F20_genomic survey sequence.
VERSION CL076017
KEYWORDS CL076017.1 GI:40531930
SOURCE GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 889)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 2
High quality sequence stop: 847.
FEATURES
Location/Qualifiers
1..889
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-138F20"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 11.5%; Score 882.6; DB 9; Length 889;
Best Local Similarity 99.6%; Pred. No. 1.6e-240;
Matches 885; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6645 TCAGAGTAAAGTTGGCCGAGTGTTATCACTCATGTTATGCGAGCAGCTGATATCTC 6704
Db 889 TCAAAAGTAGTTTGGCCGAGTGTTATCACTCATGTTATGCGAGCAGCTGATATCTC 830

Qy 6705 TTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAAGTCAACCAAGTCAT 6764
Db 829 TTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAAGTCAACCAAGTCAT 770

Qy 6765 TCTGAGAAATAGTATGCGGCGACCGAGTGTCTCTTGCCTCCGCTCAATACGGGATAATA 6824
Db 769 TCTGAGAAATAGTATGCGGCGACCGAGTGTCTCTTGCCTCCGCTCAATACGGGATAATA 710

Qy 6825 CCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTTGAAAGAGTTCTTTCGGGGCGAA 6884
Db 709 CCGCGCCACATAGCAGAACTTTAAAGTCTCATCTTTGAAAGAGTTCTTTCGGGGCGAA 650

Qy 6885 AACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATTAACCCACTCGTGCACCCA 6944
Db 649 AACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATTAACCCACTCGTGCACCCA 590

Qy 6945 ACTGATCTTACGATCTTTTACTCTTCCACGCGTTTCTGGGTGAGCAAAAAACGGAAGGC 7004
Db 589 ACTGATCTTACGATCTTTTACTCTTCCACGCGTTTCTGGGTGAGCAAAAAACGGAAGGC 530
```

```
Qy 7005 AAAATGCGCAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATCTCTTCC 7064
Db 529 AAAATGCGCAAAAGGGAATAAGGCGACACGGAATGTTGAATACTCATCTCTTCC 470
Qy 7065 TTTTTCATATTATTGAAGCAATTTATCAGGGTTATGTTCTCATGAGCGGTACATATTG 7124
Db 469 TTTTTCATATTATTGAAGCAATTTATCAGGGTTATGTTCTCATGAGCGGTACATATTG 410
Qy 7125 AATGATTTTGAAGAAATAAACAATAGGGGTTCCGCGCAATTTCCCGAAAGTGCCAC 7184
Db 409 AATGATTTTGAAGAAATAAACAATAGGGGTTCCGCGCAATTTCCCGAAAGTGCCAC 350
Qy 7185 CTGACGCTTAAGAAACCATTTATATCATGACATTTAACTATAAATAAGCGGTATCACGA 7244
Db 349 CTGACGCTTAAGAAACCATTTATATCATGACATTTAACTATAAATAAGCGGTATCACGA 290
Qy 7245 GGCCTTTTCGCTCTCGCGGTTTCGGTGATGACGCGTGAACCTCTGACATGCGAGTCC 7304
Db 289 GGCCTTTTCGCTCTCGCGGTTTCGGTGATGACGCGTGAACCTCTGACATGCGAGTCC 230
Qy 7305 CGGAGACGGTCTCACAGCTTGTCTGAAGCGGATGCGGAGCAGACAAGCCGTCAGGGCG 7364
Db 229 CGGAGACGGTCTCACAGCTTGTCTGAAGCGGATGCGGAGCAGACAAGCCGTCAGGGCG 170
Qy 7365 CGTCAGCGGGTGTGGCGGGTGTTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTG 7424
Db 169 CGTCAGCGGGTGTGGCGGGTGTTCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTG 110
Qy 7425 TACTGAGAGTGACCATATGCGGTGTGAATACCGGACAGATGCGTGAAGGAGAAATACC 7484
Db 109 TACTGAGAGTGACCATATGCGGTGTGAATACCGGACAGATGCGTGAAGGAGAAATACC 50
Qy 7485 GCATCAGCGGCATTCGCCATTCAGGCTGCGCACTGTGGGAAGGGCG 7533
Db 49 GCATCAGCGGCATTCGCCATTCAGGCTGCGCACTGTGGGAAGGGCG 1
```

RESULT 13

LOCUS CL021194/c 1028 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-8A16_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-8A16,
genomic survey sequence.

ACCESSION CL021194

VERSION CL021194.1 GI:40463007

KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 1028)

Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E., and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 175000 Std Error: 0.00

Seq primer: Sp6 ATTTAGGTGACATAATAG

Class: BAC ends

High quality sequence start: 64

High quality sequence stop: 844.

Location/Qualifiers

1..1028

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/strain="Nigerian frog"

/db_xref="taxon:9364"

/clones="CH216-8A16"

/sex="male"

FEATURES

source

```
ORIGIN
Query Match 11.4%; Score 875.6; DB 9; Length 1028;
Best Local Similarity 96.1%; Pred. No. 1.7e-238;
Matches 930; Conservative 0; Mismatches 34; Indels 4; Gaps 3;

Qy 6684 TGCAGCAGTCATATAATCTCTTACTGTCTGTCATGCGCATCCGTAAGATGCTTTTCTGTGACTG 6743
Db 978 TGCAGCAAAATCTTCTTACTGTTTCATGCGCATCCGTAAGATGCT-TTTTTCAGAACTG 920
Qy 6744 GTGAGTACTCAACCAAGTCAATCTCTGAGAATAGTGTATGCGCGCACCGAGTGTCTTTGCC 6803
Db 919 GTGAGTACTCAACCAAGTCAATCTCTGAGNATAGTGTATGCGCGCACCGAGTGTCTTTGCC 860
Qy 6804 CGCGGTCAATACGGGATAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATATTG 6863
Db 859 CGCGGTCAATACGGGATAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATATTG 800
Qy 6864 GAAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCA 6923
Db 799 GATAACGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCA 740
Qy 6924 TGTAAACCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACTTTTCAACGCGTTTCTG 6983
Db 739 TGTAAACCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACTTTTCAACGCGTTTCTG 680
Qy 6984 GGTGAGCAAAACAGGAAGCAAAATGCGCAAAAGGGAATAAGGCGCACCGAAAT 7043
Db 679 GGTGAGCAAAACAGGAAGCAAAATGCGCAAAAGGGAATAAGGCGCACCGAAAT 620
Qy 7044 GTTGAATACTCATACTCTTCTTTTCAATATTTATTGAAGCATTTATCAGGGTTATTGTC 7103
Db 619 GTTGAATACTCATACTCTTCTTTTCAATATTTATTGAAGCATTTATCAGGGTTATTGTC 560
Qy 7104 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAATAAGAGGGTTCCGCGCA 7163
Db 559 TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAATAAGAGGGTTCCGCGCA 500
Qy 7164 CATTTCCCGAAAGTGCCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCT 7223
Db 499 CATTTCCCGAAAGTGCCACCTGACGCTCTAAGAAACCATTTATCATGACATTAACCT 440
Qy 7224 ATAAAAATAGCGGTATCAGAGGCCCTTTGCTCTCGGGCGTTTCGGTGATGACGGTCAAA 7283
Db 439 ATAAAAATAGCGGTATCAGAGGCCCTTTGCTCTCGGGCGTTTCGGTGATGACGGTCAAA 380
Qy 7284 ACCTCTGACACATGACGCTCCCGAGACGCTCACAGCTTCTGTAGAGCGGATGCCGGGA 7343
Db 379 ACCTCTGACACATGACGCTCCCGAGACGCTCACAGCTTCTGTAGAGCGGATGCCGGGA 320
Qy 7344 GCAGACAAGCCGTCAGGGCGCTCAGCGGGTGTGCGGGTGTGCGGGTGTGGCTTAACT 7403
Db 319 GCAGACAAGCCGTCAGGGCGCTCAGCGGGTGTGCGGGTGTGCGGGTGTGGCTTAACT 260
Qy 7404 ATCGCGCATCAGACGAGATTGTACTAGAGTGACCATATTCGGGTGTGAAATACCGCAC 7463
Db 259 ATCGCGCATCAGACGAGATTGTACTAGAGTGACCATATTCGGGTGTGAAATACCGCAC 200
Qy 7464 GATGCGTNAAGGAGAAATACCGCATCAGGGCCATTCGCCATTCAGGCTCGCAACTGTT 7523
Db 199 GATGCGTNAAGGAGAAATACCGCATCAGGGCCATTCGCCATTCAGGCTCGCAACTGTT 140
Qy 7524 GCGAAGGGCGATCGGTGCGGGCTCTTTCGCTATTATACGCCAGCTGGCGAAAGGGGATGTG 7583
Db 139 GCGAAGGGCGATCGGTGCGGGCTCTTTCGCTATTATACGCCAGCTGGCGAAATGGGATGTG 80
Qy 7584 CTGCAAGGGCATTAAGTTGGTTAAAGCGCAGGGTTTCCAGTGTGTTGTTAAACGA 7643
Db 79 CTGCAAGGGCATTAAGAT-GGTAAACCAAGGGTTTCCAGTGTGTTGTTAAACGA 23
```

/cell_line="Stock 248 F7A2, inbred N7"

/clone_lib="CH216"

/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis

BAC library"

```
QY 7644 CGGCCAGT 7651
Db 22 CGGCCATT 15

RESULT 14
CB686151/c
LOCUS
DEFINITION
  Bn01b_02008 A
  Bn01b_AAPC_ECORC transgenic Brassica napus overexpressing BNCBF17 c
  onstitutively frost_tolerant Brassica napus cDNA clone Bn01b_02008,
  mRNA sequence.
ACCESSION
  CB686151
VERSION
  CB686151.1
KEYWORDS
  EST.
SOURCE
  Brassica napus (rape)
ORGANISM
  Brassica napus
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
  1 (bases 1 to 925)
  Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
  Chagnon, J., Farah, S., Couroux, P. and Hattori, J.
  Expressed Sequence Tags from Constitutively frost tolerant
  transgenic Brassica napus overexpressing BNCBF17
  Unpublished (2002)
  Contact: Singh, J.A.
  Eastern Cereal and Oilseed Research Centre
  Agriculture and Agri-food Canada
  KW Neaby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
  0C6, Canada
  Tel: (613) 759-1662
  Fax: (613) 759-1701
  Email: singhja@agr.gc.ca.
FEATURES
  source
    1..925
    /organism="Brassica napus"
    /mol_type="mRNA"
    /cultivar="Westar"
    /db_xref="taxon:3708"
    /clone="Bn01b_02008"
    /tissue_type="fourth leaf"
    /dev_stage="3 weeks seedling grown at room temperature"
    /clone_lib="Bn01b_AAPC_ECORC transgenic Brassica napus_ove
    repressing BNCBF17 constitutively frost_tolerant"
    /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
    Site 2: XhoI; Germinated in soil flats and seedlings grown
    for 3 weeks in a Conviron E-15 cabinet set at 20°C /16 hr
    light (250 Em-2sec-1) and 16 °C / 8 hr dark. Fourth leaves
    collected at 9 am and immediately frozen."

ORIGIN
  Query Match 11.4%; Score 871.4; DB 6; Length 925;
  Best Local Similarity 99.0%; Pred. No. 2.6e-237;
  Matches 880; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

QY 5026 CTAATCATGTCATAGCTGTTTCTGTGTGAATTTATCCGCTCAATTCACACA 5085
Db 888 CTAATCATGTCATAGCTGTTT-CTGTGTGAATTTATCCGCTCAATTCACACA 830

QY 5086 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGCCTTAAGTACGAGCTAACTCA 5145
Db 829 ACATACGAGCGGAGCATAAAGTCTAAAGCTGGGTGCCTTAAGTACGAGCTAACTCA 770

QY 5146 CATTAAATTCGTTGGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 5205
Db 769 CATTAAATTCGTTGGCTCACTGCCCGCTTTTCCAGTCGGGAAACCTGTCGCCAGCTGC 710

QY 5206 ATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTTCGGTATTGGGCGCTCTCCGCTT 5265
Db 709 ATTAATGAATCGGCCAAACCGCGGGGAGAGCGGTTTTCGGTATTGGGCGCTCTCCGCTT 650
```

```
5266 CCTCGCTCACTGACTCGCTCGCTCGGTCGGTTCGGCTCGCGGAGCGGTATCAGCTCACT 5325
Db 649 CCTCGCTCACTGACTCGCTCGCTCGGTCGGTTCGGCTCGCGGAGCGGTATCAGCTCACT 590

QY 5326 CAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAGGAAAGAACATGTGAG 5385
Db 589 CAAAGGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGCAGGAAAGAACATGTGAG 530

QY 5386 CAAAGGCGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTCGTCGCGTTTTCCTATA 5445
Db 529 CAAAGGCGCAAAAGGCGCAGGAAACCGTAAAGGCGCGTTCGTCGCGTTTTCCTATA 470

QY 5446 GGCTCCGCCCCCTTGACGAGCATCACAAAATCAAGTCAAGTCAGAGGTGGCGAAACC 5505
Db 469 GGCTCCGCCCCCTTGACGAGCATCACAAAATCAAGTCAAGTCAGAGGTGGCGAAACC 410

QY 5506 CGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAAAGCTCCTCTCGGCTCTCTG 5565
Db 409 CGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAAAGCTCCTCTCGGCTCTCTG 350

QY 5566 TTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGC 5625
Db 349 TTCCGACCTCGCGCTTACCGGATACCTGTCGGCTTCTCCCTCGGGAAGCGTGGCGC 290

QY 5626 TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAGCTGG 5685
Db 289 TTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTCGCTCCAAGCTGG 230

QY 5686 GCTGTGTGCAGAAACCCCGCTTTCAGCCGACCGCTGCGCTTATTCGGTAACTATCGTC 5745
Db 229 GCTGTGTGCAGAAACCCCGCTTTCAGCCGACCGCTGCGCTTATTCGGTAACTATCGTC 170

QY 5746 TTGAGTCCAAACCCCGTAAAGACAGCACTTATCGCACCTGGCAGCAGCCACTGGTAACAGA 5805
Db 169 TTGAGTCCAAACCCCGTAAAGACAGCACTTATCGCACCTGGCAGCAGCCACTGGTAACAGA 110

QY 5806 TTAGCAGAGCGAGGTATGTAGCGGTCTACAGAGTTCCTGAAGTGGTGGCTAACTACG 5865
Db 109 TTAGCAGAGCGAGGTATGTAGCGGTCTACAGAGTTCCTGAAGTGGTGGCTAACTACG 50

QY 5866 GCTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGT 5914
Db 49 GSTACACTAGAAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGT 1

RESULT 15
CF269652 1073 bp mRNA linear EST 13-AUG-2003
LOCUS
DEFINITION
  Fcylcolid844 Fragilariopsis cylindrus SMART cDNA library (Clontech)
  Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.
ACCESSION
  CF269652
VERSION
  CF269652.1
KEYWORDS
  EST.
ORGANISM
  Fragilariopsis cylindrus
  Fragilariopsis cylindrus
  Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
  Bacillariophycidae; Bacillariales; Bacillariaceae; Fragilariopsis.
  1 (bases 1 to 1073)
  Mock, T. and Valentin, K.
  EST analysis of freezing tolerance in the Antarctic diatom
  Fragilariopsis cylindrus: Detection of numerous cold adaption
  related genes and gene transfer events
  Unpublished (2003)
  Contact: Mock T
  Biological Oceanography
  Alfred-Wegener-Institute for Polar and Marine Research
  Am Handelshafen 12, D-27570 Bremerhaven, Germany
  Tel: +49 471 4831 1893
  Fax: +49 471 4831 1425
  Email: tmock@awi-bremerhaven.de
  sequence with unknown function
  PCR Primers
  FORWARD: 5' lambdaTriplex2
```


BACKWARD: 3'lambdaTriplex2
Seq primer: ctgaggagcgcgcattgtgttgg.
Location/Qualifiers
1. .1073
/organism="Fragilariopsis cylindrus"
/mol_type="mRNA"
/strain="Antarctic"
/db_xref="taxon:186039"
/clone_lib="Fragilariopsis cylindrus SMART cDNA library (Clontech)"
/note="vector: pTriplex2; total polyA was used for first-strand synthesis with SMART IV oligos and CDS III/3'PCR primer. Double strand cDNA synthesis was done by LD PCR using the following program: 95oC for 5 min denaturation and subsequent 20 cycles at 95oC (2min) and 68oC (6min). After SfiI digestion the cDNA was fractionated with CHROMA Spin-400 columns. These cDNAs were ligated overnight into pTriplex2 vectors."

ORIGIN

Query Match 11.3%; Score 865; DB 7; Length 1073;
Best Local Similarity 96.2%; Pred. No. 1.9e-235;
Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;
QY 5154 GCGTTCGCTCACTGCGCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGCATTATGA 5213
DB 1 GCTCTTGTCTCACTGCGCGCTTCCAGTCGGGAAACCTGTCGCCAGCTGCATTATGA 60
QY 5214 ATCGGCCAAGCGCGGGAGAGGCGTTTGGGTATTTGGGGCTCTTCGGCTTCCTCGCTC 5273
DB 61 ATCGGCCAAGCGCGGGAGAGGCGTTTGGGTATTTGGGGCTCTTCGGCTTCCTCGCTC 120
QY 5274 ACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGGGCG 5333
DB 121 ACTGACTCGCTCGCTCGCTCGCTCGCTCGCGGAGCGGTATCAGCTCACTCAAGGGCG 180
QY 5334 GTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTGAGCAAAAGGC 5393
DB 181 GTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGGAAAGAACATGTGAGCAAAAGGC 240
QY 5394 CAGCAAAAGCCAGGAAACCGTAAAGCGCGCTTGTGCTGGCGTTTTCATAGGCTCCGC 5453
DB 241 CAGCAAAAGCCAGGAAACCGTAAAGCGCGCTTGTGCTGGCGTTTTCATAGGCTCCGC 300
QY 5454 CCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGA 5513
DB 301 CCCCTCAGCAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGGA 360
QY 5514 CTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGGCTCTCCTGTTCCGACC 5573
DB 361 CTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGGCTCTCCTGTTCCGACC 420
QY 5574 CTGCGGTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGAGAGCGTGGCGTTTCTCAT 5633
DB 421 CTGCGGTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGAGAGCGTGGCGTTTCTCAT 480
QY 5634 AGCTCAGCTGTAGGTATCTCAGTTCGCTGAGTTCGCTCCAGCTGGGCTGTGTG 5693
DB 481 AGCTCAGCTGTAGGTATCTCAGTTCGCTGAGTTCGCTCCAGCTGGGCTGTGTG 540
QY 5694 CAGCAACCCCGTTACAGCCGACCGCTGGCGCTTATCCGGTAACATATCGTCTTGAGTCC 5753
DB 541 CAGCAACCCCGTTACAGCCGACCGCTGGCGCTTATCCGGTAACATATCGTCTTGAGTCC 600
QY 5754 AACCCGTAAGACACGACTTATCGCCACTGGCAGACGCACTGGTAACAGATTAGCAGA 5813
DB 601 AACCCGTAAGACACGACTTATCGCCACTGGCAGACGCACTGGTAACAGATTAGCAGA 660
QY 5814 GCGAGGTATCTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCT--AACTACGGGTACA 5871
DB 661 GCGAGGTATCTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCNTAACTACNGGCTACA 720

QY 5872 CTAGAAGGACA-GTATTTGGTATCT-GCGCTCTGCTGAAGCC---AGTTACCTTCGAAA 5926
DB 721 CTAGAAGAACANGTATTTGGTATCTNGCGCTCTGCTNGAAGCCAGNTTACGNTTCGAAA 780
QY 5927 AAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGTAGCGGTGGTTTTTGT 5986
DB 781 AAGAGTTGGTAGCTCTTGATCCGGCAACAAACCCGCTGTAGCGGTGGTTTTTGT 840
QY 5987 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC 6046
DB 841 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC 900
QY 6047 TACGGGTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGATTTTGTGATGAGATT 6106
DB 901 TACGGGTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGATTTTGGGCATGAGACA 960
QY 6107 ATCA 6110
DB 961 ATAA 964

Search completed: May 25, 2005, 19:41:54
Job time : 14663 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2005, 08:35:53 ; Search time 795 Seconds
(without alignments)
15753.537 Million cell updates/sec

Title: US-10-006-771B-1
Perfect score: 7654
Sequence: 1 agcttgcagctgcaggt.....gtaaacagcggccagtgcc 7654

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3111	40.6	5364	2	US-08-793-610-4
2	3094	40.4	6365	1	US-08-382-990-1
3	3083.4	40.3	5292	2	US-08-793-610-3
4	3083.4	40.3	5294	2	US-08-793-610-2
5	3083.4	40.3	5323	2	US-08-793-610-1
6	3082.2	40.0	6145	1	US-08-336-132-1
7	3082.2	40.0	6145	3	US-08-935-312-1
8	3082.2	40.0	6145	3	US-08-848-760B-1
9	3082.2	40.0	6145	4	US-08-826-025-1
10	3050.2	39.9	6620	4	US-08-786-531B-3
11	2703	35.3	5865	3	US-08-654-737B-3
12	2672.4	34.9	7086	4	US-09-935-194-1
13	2633.2	34.4	7560	3	US-08-844-274-20
14	2633.2	34.4	7560	4	US-08-598-421-20
15	2623.4	34.3	9164	4	US-09-205-817A-2
16	2623.4	34.3	11846	4	US-09-205-817A-4
17	2623.4	34.3	12022	4	US-09-205-817A-3
18	2623.2	34.3	7160	4	US-08-786-531B-5
19	2623	34.3	4016	3	US-09-173-053-3
20	2621.2	34.2	4950	3	US-08-789-333F-58
21	2621.2	34.2	4950	3	US-08-589-109A-7
22	2621.2	34.2	4950	3	US-08-787-738B-58
23	2621.2	34.2	4950	4	US-09-916-940-58
24	2621.2	34.2	4950	4	US-10-096-550-58
25	2620	34.2	9639	3	US-09-147-208-26
26	2620	34.2	9639	4	US-09-550-117A-26
27	2618	34.2	9632	4	US-09-403-752A-1
28	2616	34.2	4201	4	US-09-551-151A-1
29	2616	34.2	4201	4	US-09-792-568-7
30	2616	34.2	5767	4	US-09-810-861B-3
31	2616	34.2	8854	3	US-09-053-549-1
32	2615	34.2	3796	4	US-09-470-661A-32
33	2615	34.2	6295	2	US-08-659-206A-4
34	2615	34.2	7383	4	US-09-393-483A-3
35	2615	34.2	7383	4	US-09-393-483A-4
36	2614.8	34.2	4713	3	US-09-194-285-7
37	2614.8	34.2	4724	3	US-09-194-285-8
38	2612.8	34.1	4283	1	US-08-343-401A-3
39	2612.8	34.1	4283	1	US-08-445-265A-1
40	2612.8	34.1	4283	3	US-08-990-442-1
41	2612.8	34.1	4283	4	US-09-614-178-1
42	2612.8	34.1	5897	3	US-09-097-319A-26
43	2612.8	34.1	5897	4	US-09-643-971-26
44	2612.8	34.1	6898	3	US-09-097-319A-27
45	2612.8	34.1	6898	4	US-09-643-971-27

ALIGNMENTS

RESULT 1

US-08-793-610-4
; Sequence 4, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5364 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
; US-08-793-610-4

Sequence 1, Appli
Sequence 7, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 32, Appli
Sequence 4, Appli
Sequence 3, Appli
Sequence 7, Appli
Sequence 8, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 26, Appli
Sequence 26, Appli
Sequence 27, Appli
Sequence 27, Appli

Query Match		40.6%;	Score 3111;	DB 2;	Length 5364;	
Best Local Similarity		92.0%;	Pred. No. 0;	Mismatches 210;	Indels 87;	Gaps 9;
Matches 3414;		Conservative 0;				
Qy	3868	ACGAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAG	3927			
Db	1655	ACGAGCCATAGATAAAATAAAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAG	1714			
Qy	3928	ACCCACCTGTAGTTTGGCAAGC-----TAGCTTAAGTAAAGCCAT	3969			
Db	1715	ACCCACCTGTAGTTTGGCAAGCTAGAGTCGCTTAGCCTGATAGCGCAGTAAGCCAT	1774			
Qy	3970	TTTGAAGGCATGG-AAAATACATACCTCAGATAAGTTCAGATCAAGTTAGGAA	4028			
Db	1775	TTTGAAGGCATGGAAAAATACAAACCAAGTAATAGGGAATTCAGATCAAGGGCGGTA	1834			
Qy	4029	CAGAGACAGCAGATAATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTTGGCCCCGC	4088			
Db	1835	CATGAAATAGCTAACGTTTGGCCAAACAGGATATCTGCGGTGAGCAGTTTTCGGCCCCGG	1894			
Qy	4089	TCAGGGCCAAAGACAGTTTGAACAGGAGTAATATGGSCCAAAC-----AGATATCTG	4140			
Db	1895	CCGGGGCAAGAACAGATGCTCACCGCAGTTTTCGGCCCCGGCCCCGAGAACAGATG	1954			
Qy	4141	TGTTAAGCAGTTTCTCGCCCCGCTCAGGGCCAAAGACAGATGTCCTCCAGATGCGGTCCC	4200			
Db	1955	GTACCGCAGTTTTCGGCCCCGGCCCCGGGCCAAAGAACAGATGTCCTCCAGATATGCCCCA	2014			
Qy	4201	GCCTCAGCAGTTTCTAGAGAACCAATCAGATGTTTCCAGGGTGCCCCCAAGACCTGAAT	4260			
Db	2015	ACCCTCAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGGTCCCCCAAGGACCTGAAT	2074			
Qy	4261	GACCTGTGCTTATTTGAATCAACCAATCAGTTGCTTCTCGTTCGTTTCGGGGCTT	4320			
Db	2075	GACCTGTGCTTATTTGAATTAACCAATCAGCTGCTTCTCGCTTCTGTTTCGGCGCTT	2134			
Qy	4321	CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCCAGTCTCTCGA	4380			
Db	2135	CTGCTTCCCGAGCTCATAAAGAGCTCAACCCCTCACTCGGCGCCAGTCTCTCGA	2194			
Qy	4381	TAGACTGCTCGCCCGGTACCGGTGTTCTCAATAAACCCCTCTTGCAAGTTGCATCCGACT	4440			
Db	2195	TTGACTGAGTCGCGCGGTACCGGTGTTCTCAATAAACCCCTCTTGCAAGTTGCATCCGACT	2254			
Qy	4441	CGTGTCTCGCTGCTTCTTGGAGGGTCT-CTCTGAGTGAATTGAATACCCGTCAGC-GGG	4498			
Db	2255	CGTGTCTCGCTGCTTCTTGGAGGGTCTCTCTGAGTGAATTGAATACCCGTCAGCAGGGG	2314			
Qy	4499	GTCTTTCAGTTCTCCACCTACACAGTCTCACTAAACATTCCTGATGTGCGCAGGGAC	4558			
Db	2315	GTCTTTCAGTTCTCCACCTACACAGTCTCACTAAACATTCCTGATGTGCGCAGGGAC	2374			
Qy	4559	TCCGTCAGCCCGGTTTTGTTTATATAAATAAAGCAAGAACAGTGTTCCTTCAAGCCAGA	4618			
Db	2375	TCCGTCAGCCCGGTTTTGTTTATATAAATAAAGCAAGAACAGTGTTCCTTCAAGCCAGA	2434			
Qy	4619	CTACATCTGACTCTCGGCTTTATAAAGAAATTTGAAGGGCTCTGTGGAATATCTGCCA	4678			
Db	2435	CTACATCTGACTCTCGGCTTTATAAAGAAATTTGAAGGGCTCTGTGGAATATCTGCCA	2494			
Qy	4679	CACGACTTTTAAAGATTTTATGCTTCCCTCCGATGAGGATTTAGTCAATCTATCTCGTC	4738			
Db	2495	CACGAC-TTTTAAGATTTTATGCTTCCCTCCGATGAGGATTTAGTCAATCTATCTCGTC	2553			
Qy	4739	TATTTTGTGCTTCTCCGTAATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCAG	4798			
Db	2554	TATTTGTGCTTCTCCGTAATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCAG	2613			
Qy	4799	GGATTTGCAAGTAGTTAATATCTGAGGGCAGGCTTCTGTGAAAGAGTTGCTGGGCTC	4858			
Db	2614	GGATTTGCAAGTAGTTAATATCTGAGGGCAGGCTTCTGTGAAAGAGTTGCTGGGCTC	2673			

Qy	4859	AGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACAGACAGATCGGAATCTA	4918			
Db	2674	AGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACAGACAGATCGGAATCTA	2733			
Qy	4919	GAGTGATATCTCAGAGTCCCCCGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGCTTAG	4978			
Db	2734	GAGTGATATCTCAGAGTCCCCCGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGCTTAG	2793			
Qy	4979	AGATCCTCTACCGCGGACGATCGTGGCGGGGTACCGAGCTCGAATTCGTATCATGGTC	5038			
Db	2794	AGATCCTCTACCGCGGACGATCGTGGCGGGCATCACCGGCGCCACAGGTGCGGTGCTG	2853			
Qy	5039	ATAGCTGTTC-----CTGTGTGAAATTTGTATCCGCTCAAAATTCACACA	5085			
Db	2854	GGGCTTATATCGCCGACATCACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCATGA	2913			
Qy	5086	ACATACGAGCGGAAGCATAAAGTGTAAAGCTTGGGGTGCCTTAATGAGT-----	5134			
Db	2914	GGGCTTGTTCGGCGTGGGTATGTGGCAGGCCCCGTGCGCGGGGACTGTTGGGCGCCA	2973			
Qy	5135	-----GAGCTAACTCACATTAATTCGGTTGCG	5162			
Db	2974	TCTCTTCGATGCACCATTCCTTGGCGGGGGTCTCAACGGCTCAACTACTACTCGG	3033			
Qy	5163	TCACCTGCCCGCTTTCAGTCCGGGAAACCTGTCTGCCAGCTGCATTAATGAATCGGCCAA	5222			
Db	3034	GCTGCTTCTTAATGACGAGTGCATAGGGAGAGGCTCCTGCATTAATGAATCGGCCAA	3093			
Qy	5223	CGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTTCGGTTCCTCGCTCACTGACTCG	5282			
Db	3094	CGCGGGGAGAGCGGTTTGGGTATTTGGCGCTCTTCGGTTCCTCGCTCACTGACTCG	3153			
Qy	5283	CTGCGCTCGTGTTCGGCTCGCGGAGGGTATCAGCTCACTCAAGGCGGTATACGG	5342			
Db	3154	CTGCGCTCGTGTTCGGCTCGCGGAGCGGTATCAGCTCACTCAAGGCGGTATACGG	3213			
Qy	5343	TTATCCACAGATCAGGGGATAACGACAGAAAGAAACATGTGAGCAAAAGGCCAGAAAAG	5402			
Db	3214	TTATCCACAGATCAGGGGATAACGACAGAAAGAAACATGTGAGCAAAAGGCCAGAAAAG	3273			
Qy	5403	GCCAGAAACCGTAAAGGCGCGTTGTGGCGTTTTTCCATAGGCTCGCCCCCTGAC	5462			
Db	3274	GCCAGAAACCGTAAAGGCGCGTTGTGGCGTTTTTCCATAGGCTCGCCCCCTGAC	3333			
Qy	5463	GAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGA	5522			
Db	3334	GAGCATCAAAAATCGACGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGA	3393			
Qy	5523	TACCAAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGCGCTT	5582			
Db	3394	TACCAAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTGCGCTT	3453			
Qy	5583	ACCGGATACCTGTTCGCGCTTCTCCTTCGGGAAGCGGTGGCGCTTCTCATAGCTCACGC	5642			
Db	3454	ACCGGATACCTGTTCGCGCTTCTCCTTCGGGAAGCGGTGGCGCTTCTCATAGCTCACGC	3513			
Qy	5643	TGTAGTATCTCAGTTCGGTGTAGTCTCGTCCAGCTCGGGCTGTGTGACCAAGCCC	5702			
Db	3514	TGTAGTATCTCAGTTCGGTGTAGTCTCGTCCAGCTCGGGCTGTGTGACCAAGCCC	3573			
Qy	5703	CCCGTTACGCGCCGACCGCTTATCCGCTTATCCGCTTATCCGCTTTCGAGTCCAAACCGGTA	5762			
Db	3574	CCCGTTACGCGCCGACCGCTTATCCGCTTATCCGCTTATCCGCTTTCGAGTCCAAACCGGTA	3633			
Qy	5763	AGACACGATTTATCGCCACTGGCAGACCACTGTGTAAACAGATTTAGCAGAGCGAGTAT	5822			
Db	3634	AGACACGATTTATCGCCACTGGCAGACCACTGTGTAAACAGATTTAGCAGAGCGAGTAT	3693			
Qy	5823	GTAGGGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAGTGGCTTACCTAGAGAGGACA	5882			
Db	3694	GTAGGGGTGTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTTACATAGAGAGACA	3753			
Qy	5883	GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAAAAGAGTTGCTGCTCT	5942			

Db 3754 GTATTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCTTCGGAAAAAGAGTTGGTAGCTCT 3813
Qy 5943 TGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTGGTTTGGCAAGCAGCAGATT 6002
Db 3814 TGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGTTTTTTTGGTTTGGCAAGCAGCAGATT 3873
Qy 6003 ACGCGCAGAAAAAAGAGTCTCAAGAGATCTTTTGATCTTTTCTACCGGGGTCTGACGCT 6062
Db 3874 ACGCGCAGAAAAAAGAGTCTCAAGAGATCTTTTGATCTTTTCTACCGGGGTCTGACGCT 3933
Qy 6063 CAGTGGAAACGAAACTCAGCTTTAAGGGATTTTGGTTCATGAGATATCAAAAGAGATCTTC 6122
Db 3934 CAGTGGAAACGAAACTCAGCTTTAAGGGATTTTGGTTCATGAGATATCAAAAGAGATCTTC 3993
Qy 6123 ACCTAGATCTTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATGAGTAA 6182
Db 3994 ACCTAGATCTTTTAAATTAAGAGTTTAAATCAATCTAAAGTATATATGAGTAA 4053
Qy 6183 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGAGCCACCTATCTCAGCGATCTGTCTA 6242
Db 4054 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGAGCCACCTATCTCAGCGATCTGTCTA 4113
Qy 6243 TTTGGTTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAAATACCATACGGAGGGC 6302
Db 4114 TTTGGTTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAAATACCATACGGAGGGC 4173
Qy 6303 TTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGAT 6362
Db 4174 TTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCACGCTCACCGGCTCCAGAT 4233
Qy 6363 TTATCAGCAATAAACACGACGAGCGGAGGCGGAGCGAGAGTGTCTTGCACATTTA 6422
Db 4234 TTATCAGCAATAAACACGACGAGCGGAGGCGGAGCGAGAGTGTCTTGCACATTTA 4293
Qy 6423 TCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGATAGTTCGCCAGTT 6482
Db 4294 TCCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGATAGTTCGCCAGTT 4353
Qy 6483 AATAGTTTGGCAACGTTGTTGCCATTCCTACAGGC-TGCTGTTGCTCAGCTCGTGGTTT 6541
Db 4354 AATAGTTTGGCAACGTTGTTGCCATTCCTACAGGCATCTGTTGCTCAGCTCGTGGTTT 4413
Qy 6542 GGTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATG 6601
Db 4414 GGTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCCCCATG 4473
Qy 6602 TTGTGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCGTTGTCAGAGTAAGTTGGCC 6661
Db 4474 TTGTGCAAAAAGCGGTTAGCTCTTCCGCTCCGATCGTTGTCAGAGTAAGTTGGCC 4533
Qy 6662 GCAGTGTTCATCATCTGTTATGSCAGCATGTCATTAATCTTCTTACTGTCTATGCCATCC 6721
Db 4534 GCAGTGTTCATCATCTGTTATGSCAGCATGTCATTAATCTTCTTACTGTCTATGCCATCC 4593
Qy 6722 GTAAGATGCTTTTCTGCTAGCTGGTAGTCAACCAAGTCAATCTGAGAAATAGTGTATG 6781
Db 4594 GTAAGATGCTTTTCTGCTAGCTGGTAGTCAACCAAGTCAATCTGAGAAATAGTGTATG 4653
Qy 6782 CGGCGACCGAGTTGCTTTCGCGCGCTCAATACGGGATATACCGGCCACATAGCAGA 6841
Db 4654 CGGCGACCGAGTTGCTTTCGCGCGCTCAATACGGGATATACCGGCCACATAGCAGA 4713
Qy 6842 ACTTTAAAGTGTCTCATCTATTGGAAAAAGTTTCTTCGGGCGCAAACTCTCAAGATCTTA 6901
Db 4714 ACTTTAAAGTGTCTCATCTATTGGAAAAAGTTTCTTCGGGCGCAAACTCTCAAGATCTTA 4773
Qy 6902 CCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATCTTCAGCATCT 6961
Db 4774 CCGCTGTTGAGATCCAGTTCGATGTAAACCCACTCGTGACCCAACTGATCTTCAGCATCT 4833
Qy 6962 TTTACTTTTACAGGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCGCGCAAAAAAG 7021

Db 4834 TTTACTTTTACAGGTTTCTGGGTGAGCAAAAAACAGGAAGCAAAATGCCCAAAAAAG 4893
Qy 7022 GGAATAAGGGCGACACGAAATGTTGAATCTACTACTCTTCTCTTTTCAATATTATTGA 7081
Db 4894 GGAATAAGGGCGACACGAAATGTTGAATCTACTACTCTTCTCTTTTCAATATTATTGA 4953
Qy 7082 AGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGTAATGATTTAGAAAAAT 7141
Db 4954 AGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTGTAATGATTTAGAAAAAT 5013
Qy 7142 AAAACAATAGGGTTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTCTAAGAAACC 7201
Db 5014 AAAACAATAGGGTTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTCTAAGAAACC 5073
Qy 7202 ATTATTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCCCTTTCTCTCGCG 7261
Db 5074 ATTATTATCATGACATTAACCTATAAAATAGGCGTATCAGAGGCCCTTTCTCTCGCG 5133
Qy 7262 CGTTTCGGTATGAGCGTGAACCTCTGACACATGACAGCTCCCGGAGACGCTCAGCT 7321
Db 5134 CGTTTCGGTATGAGCGTGAACCTCTCTGACACATGACAGCTCCCGGAGACGCTCAGCT 5193
Qy 7322 TGTCTGTAAGCGGATCCCGGAGCAGACAAGCCGCTCAGGCGCTGAGCGGTGTTGCG 7381
Db 5194 TGTCTGTAAGCGGATCCCGGAGCAGACAAGCCGCTCAGGCGCTGAGCGGTGTTGCG 5253
Qy 7382 GGGTGTGCGGCTCGCTTAACTATCGGCGATCAGAGCAGATTGTACTGAGAGTGCACCAT 7441
Db 5254 GGGTGTGCGGCTCGCTTAACTATCGGCGATCAGAGCAGATTGTACTGAGAGTGCACCAT 5313
Qy 7442 ATGCGGTGTGAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGG 7492
Db 5314 ATGCGGTGTGAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGG 5364

RESULT 2

US-08-352-990-1
; Sequence 1, Application US/08352990
; Patent No. 5681562
; GENERAL INFORMATION:
; APPLICANT: SOBOL, ROBERT E
; APPLICANT: GAGE, FRED H
; APPLICANT: ROYSTON, IVOR
; APPLICANT: FRIEDMAN, THEODORE
; TITLE OF INVENTION: LYMPHOKINE GENE THERAPY OF CANCER
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 91211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/352,990
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/863,641
; FILING DATE: 19920403
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-SD 9295
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 1:

Db 3982 CCGCTCAGCGGGGCTTTTCAATTTGGGGCTCGTCCGGATCGGAGACCCCTGCCAGG 4041
 QY 4547 TGCCGAGGAGCTCCGCTCAGCCCGGTTTTTGTATATAAATAAGCAAGTGTTC 4606
 Db 4042 GACCACCGACCCACACCGGAGGTAAGCTG----- 4072
 QY 4607 CTTCAAGCCAGACTACATCCTCGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTG 4666
 Db 4073 -----GCTGCTCGCGGTTTCGGTGATGACGGTGAACCTCTG-A 4113
 QY 4667 GACTATCTGCACACAGACTTTTAAAGATTTTATGCTCCTCGATGAGGATTTAGTCAA 4726
 Db 4114 CACATCAGCTCCCGAGACGGTCAACAGCTTGTCTGTAAGCGGATGCCGGAGACACAA 4173
 QY 4727 TCTATCCTCGTCTATTTTCTGCTTCTCGTCTCTCGTATTTTAAATTTCTAGTTTGCACCTT 4786
 Db 4174 GCCCGTCAGGGCGGTCAGCGGGTGTGGCG----- 4204
 QY 4787 CCTGAGAGCAGCGGATTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAGG 4846
 Db 4205 ----- 4204
 QY 4847 TTGCTGGGCTCAGTGTGAGATTTTCCCATAAAGGGTCTCTGCCCTGTGTACAGACA 4906
 Db 4205 ----- 4204
 QY 4907 GATCGGAATCAGATGTCATCTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGC 4966
 Db 4205 -----GGTCTCGGGGCGCAGCCATGACCCA 4229
 QY 4967 ATCTTTTGCCTAGAGATCCTCTACCGCGAGCGATCGTGCGGCTACCGAGCTCGAATTC 5026
 Db 4230 GTC-----ACGTAGCATAGCGGAGTGATATCTGGCTTAA 4264
 QY 5027 GTAATCATGTCATAGCTGTTCTGTGTGAATTTGTTATCCGCTCACAAITCCACAAA 5086
 Db 4265 CTATGGGTCATCAGACGATTTGACTGAGAG----- 4296
 QY 5087 CATACAGCGGAGCATAAAGTGTAAAGCTGGGGTGGCTTAATGAGTGAGTAACTCAC 5146
 Db 4297 ----- 4296
 QY 5147 ATTAATTTGGCTTGGCTCACTGCCGCTTTCAGTCGGGAACCTGTGTCGACGTGCA 5206
 Db 4297 -----TGCACCATATGGC 4309
 QY 5207 TTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTTCGTAATGGCGCTCTTCCGCTTC 5266
 Db 4310 GTGTGAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTC 4369
 QY 5267 CTCGCTCACTGACTCGCTGGCTCGTTCGCTCGGCTGCGGAGCGGATCAGCTCACTC 5326
 Db 4370 CTCGCTCACTGACTCGCTCGCTCGTTCGCTCGGCTGCGGCGAGCGGATCAGCTCACTC 4429
 QY 5327 AAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGAGGAAGAAACATGTGAGC 5386
 Db 4430 AAAGCGGTAAATACGGTTATCCAGAAATCAGGGGATAACGAGGAAGAAACATGTGAGC 4489
 QY 5387 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGCGGTTGCTGGCGTTTTTCCATAG 5446
 Db 4490 AAAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGCGGTTGCTGGCGTTTTTCCATAG 4549
 QY 5447 GCTCCGCCCTTGACGAGCATCAAAAATGACGCTCAAGTCAGAGTGGCGAAACCC 5506
 Db 4550 GCTCCGCCCTTGACGAGCATCAAAAATGACGCTCAAGTCAGAGTGGCGAAACCC 4609
 QY 5507 GACGAGCTATAAAGATACAGCGCTTCCCTGGAAGCTCCCTCGCTCGCTCTCTCTGT 5566
 Db 4610 GACGAGCTATAAAGATACAGCGGTTTCCCTGGAAGCTCCCTCGCTCGCTCTCTCTGT 4669
 QY 5567 TCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGGGAAGCGTGGCGT 5626
 Db 4670 TCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTTCGGGAAGCGTGGCGT 4729

QY 5627 TTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGTCGAAGCTGG 5686
 Db 4730 TTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGTTAGGTGCTTCCGTCGAAGCTGG 4789
 QY 5687 CTGTGTGACGAAACCCCGTTTACGCCGACCGCTTATCCGGTAACTATCGTCT 5746
 Db 4790 CTGTGTGACGAAACCCCGTTTACGCCGACCGCTTATCCGGTAACTATCGTCT 4849
 QY 5747 TGAGTCCAAACCCGGTAAACACGACTTATCGCACTGGCAGCAGCACTGGTAAACAGAT 5806
 Db 4850 TGAGTCCAAACCCGGTAAACACGACTTATCGCACTGGCAGCAGCACTGGTAAACAGAT 4909
 QY 5807 TAGCAGAGCAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGG 5866
 Db 4910 TAGCAGAGCAGGTATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCCTAACTACGG 4969
 QY 5867 CTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGTGAAAGCAGTTACCTTCGAAA 5926
 Db 4970 CTACACTAGAAGCAGATTTTGGTATCTGCGCTCTGTGAAAGCAGTTACCTTCGAAA 5029
 QY 5927 AAGAGTTGGTAGTCTTGTATCCGGCAAAACAAACCGCTGTGAGGGTGGTTTTTGT 5986
 Db 5030 AAGAGTTGGTAGTCTTGTATCCGGCAAAACAAACCGCTGTGAGGGTGGTTTTTGT 5089
 QY 5987 TTGCAAGCAGCAGATTTACGCGCAGAAAAAGGATCTCAAGAGATCTTTCGATCTTTTC 6046
 Db 5090 TTGCAAGCAGCAGATTTACGCGCAGAAAAAGGATCTCAAGAGATCTTTCGATCTTTTC 5149
 QY 6047 TACGGGCTCTCAGCTCAGTGGAAACGAACTCAGCTTAAAGGATTTTGGTCATGAGATT 6106
 Db 5150 TACGGGCTCTCAGCTCAGTGGAAACGAACTCAGCTTAAAGGATTTTGGTCATGAGATT 5209
 QY 6107 ATCAAAAAGGATCTTCACTAGATCTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6166
 Db 5210 ATCAAAAAGGATCTTCACTAGATCTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 5269
 QY 6167 AAGTATATAGTAAACCTTGTCTGACAGTTTCAATGCTTAAATCAGTGAGGCACTTAT 6226
 Db 5270 AAGTATATAGTAAACCTTGTCTGACAGTTTCAATGCTTAAATCAGTGAGGCACTTAT 5329
 QY 6227 CTCAGGATCTGCTATTTTCTCATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6286
 Db 5330 CTCAGGATCTGCTATTTTCTCATCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5389
 QY 6287 TACGATAGCGGAGGCTTACATCTGCGCCAGAGTGTGCAATGATACCGCAGAGACCGAG 6346
 Db 5390 TACGATAGCGGAGGCTTACATCTGCGCCAGAGTGTGCAATGATACCGCAGAGACCGAG 5449
 QY 6347 CTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGCGAGAG 6406
 Db 5450 CTCACCGGCTCCAGATTTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGCGAGAG 5509
 QY 6407 TGGTCTGCAACTTTTATCGGCTTCAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 6466
 Db 5510 TGGTCTGCAACTTTTATCGGCTTCAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 5569
 QY 6467 AAGTATAGTGTGCGGATTAATGTTGCGCAACGTTGTTGCAATTTGTTGCAATTTGTTG 6525
 Db 5570 AAGTATAGTGTGCGGATTAATGTTGCGCAACGTTGTTGCAATTTGTTGCAATTTGTTG 5629
 QY 6526 GTACAGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 6585
 Db 5630 GTACAGCTCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5689
 QY 6586 TACATGATCCCGCATTTGTTGCAAAAAAGCGGTTAGCTCTTTCGCTCTCCGATCGTTGT 6645
 Db 5690 TACATGATCCCGCATTTGTTGCAAAAAAGCGGTTAGCTCTTTCGCTCTCCGATCGTTGT 5749
 QY 6646 CAGAAATAGTGTGCGCAGTGTATCACTCATGTTTATGGCAGCACTGTCATAATTTCTCT 6705
 Db 5750 CAGAAATAGTGTGCGCAGTGTATCACTCATGTTTATGGCAGCACTGTCATAATTTCTCT 5809

```
QY 6706 TACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATT 6765
Db 5810 TACTGTCTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCAATT 5869
QY 6766 CTGAGAAATAGTGTATGCGCGCAGCAGAGTTGCTCTTGGCCCGCGTCAATACGGGATAATAC 6825
Db 5870 CTGAGAAATAGTGTATGCGCGCAGCAGAGTTGCTCTTGGCCCGCGTCAACACGGGATAATAC 5929
QY 6826 CGCGCCACATGACAGAACTTTAAAGTGTCTATCTTGGGAAACGTTCTTGGGGCGGAAA 6885
Db 5930 CGCGCCACATGACAGAACTTTAAAGTGTCTATCTTGGGAAACGTTCTTGGGGCGGAAA 5989
QY 6886 ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGATTCAGATGTAACCCACTCGTGACCCCAA 6945
Db 5990 ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGATTCAGATGTAACCCACTCGTGACCCCAA 6049
QY 6946 CTGATCTTTCAGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCA 7005
Db 6050 CTGATCTTTCAGCATCTTTTACTTTTCCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCA 6109
QY 7006 AAATGCCGCAAAAAGGAATAAGGGCGACAGCGAAATGTTGAATACATCACTTCTTCCCT 7065
Db 6110 AAATGCCGCAAAAAGGAATAAGGGCGACAGCGAAATGTTGAATACATCACTTCTTCCCT 6169
QY 7066 TTTTCAATATTATTGAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATATTTGA 7125
Db 6170 TTTTCAATATTATTGAAGCATTTATCAGGTTTATGCTCTCATGAGCGGATACATATTTGA 6229
QY 7126 ATGATTATTAGAAAATAAACAATAGGGTTCGGCGCATTTCCCGGAAAGTGCCACC 7185
Db 6230 ATGATTATTAGAAAATAAACAATAGGGTTCGGCGCATTTCCCGGAAAGTGCCACC 6289
QY 7186 TGAGCTCTAAGAAACCATTTATCATGACATTAACCTTATAAATAATAGGCGGTATCACGAG 7245
Db 6290 TGAGCTCTAAGAAACCATTTATCATGACATTAACCTTATAAATAATAGGCGGTATCACGAG 6349
QY 7246 GCCCTTTCGTCT 7257
Db 6350 GCCCTTTCGTCT 6361
```

RESULT 3

```
US-08-793-610-3
; Sequence 3, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5292 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-3
```

Query Match 40.3%; Score 3083.4; DB 2; Length 5292;

Best Local Similarity 91.8%; Pred. No. 0;
Matches 3400; Conservative 0; Mismatches 191; Indels 112; Gaps 9;

```
QY 3868 ACGAGCCATAGATAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAG 3927
Db 1624 ACGAGCCATAGATAAATAAAGATTTTATTAGTCTCCAGAAAAGGGGGAATGAAG 1683
QY 3928 ACCCCACCTGTAGGTTTGGCAAGC-----TAGCTTAAAGTAACGCCAT 3969
Db 1684 ACCCCACCTGTAGGTTTGGCAAGCTAGAGTCGCTTAGCTGATAGCCGAGTAAGCCAT 1743
QY 3970 TTTCAAGGCGATG-AAAATAATACATGAGATAGAGAAGTTTCAAGTTAGGTTAGAA 4028
Db 1744 TTTCAAGGCGATGAAAATAATACCAAAACCAAGATAGGGAAGTTTCAAGTCAAGGGCGGTA 1803
QY 4029 CAGAGACACGACAAATATGSGCCAAACAGAGATATCTGTGTAGAGAGTTTCTGCCCCGC 4088
Db 1804 CATGAAATAGCTAAACGTTGGGCCAAACAGAGATATCTGCGGTGAGCAGTTTCGGCCCCGG 1863
QY 4089 TCAGGGCCAAAGACAGTTTGGAAACAGGAGATATGGGCCAAACAGAGATATCTGTGTGAAGC 4148
Db 1864 CCCGGGCGNAGACAGATG-----TCACGCG 1890
QY 4149 AGTTCTGCCCCGCTCAGGGCCAAAGAACAGATGTTCCCAAGATCGCGTCCCGCCCTCAG 4208
Db 1891 AGTTTCGCCCCGCGCCGAGGCCAAAGAACAGATGTTCCCAAGATATGATGCCCAACCTCAG 1950
QY 4209 CAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCCTGT 4268
Db 1951 CAGTTTCTTAAAGACCCATCAGATGTTTCCAGGCTCCCCCAAGGACCTGAAATGACCCTGC 2010
QY 4269 GCCTTATTGAACTAAACCAATCAGTTCGCTTCTGCTTCTGTCGCGGCTTCTGCTCCC 4328
Db 2011 GCCTTATTGAAATTAACAATCAGCTGCTTCTGCTTCTGTCGCGGCTTCTGCTTCC 2070
QY 4329 CGAGCTCAATAAAGAGCCCAACAAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGC 4388
Db 2071 CGAGCTCTATAAAGAGCTCACAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGA 2130
QY 4389 GTGCGCGGGTACCCGTGTTCTCAATAAACCCCTTCTGAGTTGATCCGACTCGTGTCT 4448
Db 2131 GTGCGCGGGGTACCCGTGTTCTCAATAAACCCCTTCTGAGTTGATCCGACTCGTGTCT 2190
QY 4449 CGCTGTTCTTGGGAGGCTCT-CTCTGAGTGAATGACTACCCGTCAGC-GGGGTCTTTCA 4506
Db 2191 CGCTGTTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCCGTCAGGCGGGGCTTTTCA 2250
QY 4507 GTTTCTCCCACTACACAGGCTCTCAATAACATTTCTGATGTGCGGAGGAGTCCGTCAG 4566
Db 2251 GTTTCTCCCACTACACAGGCTCTCAATAACATTTCTGATGTGCGGAGGAGTCCGTCAG 2310
QY 4567 CCGGTTTTTGTATTATAAATAAAGAACAGTGTTCCTTCAAGCCAGACTACATCC 4626
Db 4626 CCGGTTTTTGTATTATAAATAAAGAACAGTGTTCCTTCAAGCCAGACTACATCC 4626
```

Db 2311 CCGGTTTGTGTTTAAATAAATGCAAGACAGTGTTCCTTTCAAGCCAGACTACATCC 2370
QY TGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCCACACAGCTT 4686
Db 2371 TGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCCACACAGC-T 2429
QY TTTAAGATTTTATGCTCTCTGGATGAGGAAATTAAGTCAATCTATCTCTGCTATTTTGC 4746
Db 2430 TTTAAGATTTTATGCTCTCTGGATGAGGAAATTAAGTCAATCTATCTCTGCTATTTTGC 2489
QY TGGCTTCTCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCACGGCGATTGC 4806
Db 2490 TGGCTTCTCGTATTTTAAATTTCTAGTTTGCACCTCCCTTCTGAGAGCACGGCGATTGC 2549
QY AGAGTAGTTAAATCTCTGAGGCGAGGCTTCTGTGAAAAGTTTGCCTGGGCTCAGTGTGAG 4866
Db 2550 AGAGTAGTTAAATCTCTGAGGCGAGGCTTCTGTGAAAAGTTTGCCTGGGCTCAGTGTGAG 2609
QY ATTTTSCCATAAAAAGGGTCTTGCCCTCTGTGTACAGACAGATCGGAATCTAGAGTGCAT 4926
Db 2610 ATTTTSCCATAAAAAGGGTCTTGCCCTCTGTGTACAGACAGATCGGAATCTAGAGTGCAT 2669
QY ACTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGGCTAGAGATCCCTC 4986
Db 2670 ACTCAGAGTCCCGCGGTTCCGGGCTCTGATCTCAGGGCATCTTTGGCTAGAGATCCCTC 2729
QY TAGCGCGAGCATCTGCGCGGTACCGAGCTCGAATTCGTAATCATGTCATAGCTGT 5046
Db 2730 TAGCGCGAGCATCTGCGCGGTACCGAGCTCGAATTCGTAATCATGTCATAGCTGT 2789
QY TTC- - - - -CTGTGTGAATTTGTATTCGCTCACAATTCACAACAATACGA 5093
Db 2790 ATCGCGACATCACCGATGGGAGATCGGGTCGCCACTTCGGGCTCATGAGCGTGT 2849
QY GCGGGAAGCATAAAGTGAAGCTGGGTGCTTAATGAT- - - - - 5134
Db 2850 TTGCGGCTGGTATGTGGCAGCCCGGTGGCGGGGACTGTGGGCGCATCTCCTTG 2909
QY 5135 - - - - -GAGCTAACTCACAATTAATGCTTGGCTCAGTGC 5170
Db 2910 CATGACCAATTCCTGGCGGCGGTGCTCAAGCGCTCAACCTACTACTGGGCTGCTTC 2969
QY 5171 CGCTTTCCAGTGGGAAACCTGTCTGCCAGCTGCAATTAATGAATCGGCAACGGCGGG 5230
Db 2970 CTAACTCAGAGTGCATAAAGGAGAGCGTCTGCAATTAATGAATCGGCAACGGCGGG 3029
QY 5231 GAGAGCGGTTTGGGCTTTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 5290
Db 3030 GAGAGCGGTTTGGGCTTTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 3089
QY 5291 GGTGTTCCGCTGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGTTATCCAC 5350
Db 3090 GGTGTTCCGCTGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGTTATCCAC 3149
QY 5351 AGAATCAGGGATAACGAGGAAAGAAATGTGAGCAAAAGCGCAGCAAAAGCCAGGAA 5410
Db 3150 AGAATCAGGGATAACGAGGAAAGAAATGTGAGCAAAAGCGCAGCAAAAGCCAGGAA 3209
QY 5411 CCGTAAAAAGCGCGTGTCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCA 5470
Db 3210 CCGTAAAAAGCGCGTGTCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCA 3269
QY 5471 CAAAATCAGCGTCAAGTCAGAGTGGCGAAACCGCAGAGCTATAAAGATACGAGC 5530
Db 3270 CAAAATCAGCGTCAAGTCAGAGTGGCGAAACCGCAGAGCTATAAAGATACGAGC 3329
QY 5531 GTTTCCTCCGTAAGCTCCCTGTGGCTCTCTGTTCGACCCCTGCGCTTACCGGATA 5590
Db 3330 GTTTCCTCCGTAAGCTCCCTGTGGCTCTCTGTTCGACCCCTGCGCTTACCGGATA 3389
QY 5591 CTGTCCGCTTCTCCCTTCGGGAAGCGGTGCGCTTCTCATAGCTCACTGATAGTA 5650
Db 3390 CTGTCCGCTTCTCCCTTCGGGAAGCGGTGCGCTTCTCATAGCTCACTGATAGTA 3449

QY 5651 TCTCAGTTCCGTTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTCA 5710
Db 3450 TCTCAGTTCCGTTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTCA 3509
QY 5711 GCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGCGTAAAGACGA 5770
Db 3510 GCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACCGCGTAAAGACGA 3569
QY 5771 CTTATCGCCACTGCGCAGCAGCACCTGTAACAGGATTAAGCAGAGCGAGTATGTAGGCG 5830
Db 3570 CTTATCGCCACTGCGCAGCAGCACCTGTAACAGGATTAAGCAGAGCGAGTATGTAGGCG 3629
QY 5831 TGCTACAGAGTTCTTGAAGTGGTCCCTAACTTACGGCTACACTAGAAGACAGTATTTGG 5890
Db 3630 TGCTACAGAGTTCTTGAAGTGGTCCCTAACTTACGGCTACACTAGAAGACAGTATTTGG 3689
QY 5891 TATTCGCTCTGTGAAAGCCAGTTACCTTCGGAAGAAAGAGTTGTAGCTCTTTGATCCGG 5950
Db 3690 TATTCGCTCTGTGAAAGCCAGTTACCTTCGGAAGAAAGAGTTGTAGCTCTTTGATCCGG 3749
QY 5951 CAAACAAACCCCGCTGGTAGCGGTGTTTTTGTTCGAGCAGCAGATTAACGGCAG 6010
Db 3750 CAAACAAACCCCGCTGGTAGCGGTGTTTTTGTTCGAGCAGCAGATTAACGGCAG 3809
QY 6011 AAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAA 6070
Db 3810 AAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGGAA 3869
QY 6071 CGAAAACTCAGCTTAAGGATTTTGGTCAAGATTAATCAAAAAAGATCTTCCACCTAGAT 6130
Db 3870 CGAAAACTCAGCTTAAGGATTTTGGTCAAGATTAATCAAAAAAGATCTTCCACCTAGAT 3929
QY 6131 CCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGATATATATGATAGTAACCTTGTTC 6190
Db 3930 CCTTTTAAATTAATAATGAAGTTTAAATCAATCTAAAGATATATATGATAGTAACCTTGTTC 3989
QY 6191 TGACAGTTTACCAATCTTAATCAGTGAGCAGCTATCTCAGCGATCTGTCTATTTTCGTTTC 6250
Db 3990 TGACAGTTTACCAATCTTAATCAGTGAGCAGCTATCTCAGCGATCTGTCTATTTTCGTTTC 4049
QY 6251 ATCCATAGTTGCTGACTCCCCGCTCGTGTAGATAAATAAGATACGAGGAGGCTTACCATC 6310
Db 4050 ATCCATAGTTGCTGACTCCCCGCTCGTGTAGATAAATAAGATACGAGGAGGCTTACCATC 4109
QY 6311 TGGCCCCAGTGTGCAATGATACCGCAGACCCAGCCTCACCGGCTCCAGATTTATCAGC 6370
Db 4110 TGGCCCCAGTGTGCAATGATACCGCAGACCCAGCCTCACCGGCTCCAGATTTATCAGC 4169
QY 6371 AATAAACCCAGCAGCGGAAGCGGCGAGCAGAGTGTCTCTCAACTTTATCCGCTTC 6430
Db 4170 AATAAACCCAGCAGCGGAAGCGGCGAGCAGAGTGTCTCTCAACTTTATCCGCTTC 4229
QY 6431 CATCCAGTCTATTAATTTGTTCGCGGAAGCTAGAGTAAAGTGTTCGCCAGTTAATAGTTT 6490
Db 4230 CATCCAGTCTATTAATTTGTTCGCGGAAGCTAGAGTAAAGTGTTCGCCAGTTAATAGTTT 4289
QY 6491 GCGCAACGTTGTGCAATGTCTACAGGC-TGTGTGTGTCACGCTCGTGTGTTGTTATGCG 6549
Db 4290 GCGCAACGTTGTGCAATGTCTACAGGCATCGTGTGTGTCACGCTCGTGTGTTGTTATGCG 4349
QY 6550 TTTCATTACGCTCCGCTTCCCAACGATCAAGGAGGTTACATGATCCCCCATGTTGTGCAA 6609
Db 4350 TTTCATTACGCTCCGCTTCCCAACGATCAAGGAGGTTACATGATCCCCCATGTTGTGCAA 4409
QY 6610 AAAAGCGTTAGCTCCTTTCGGTCTCTCCGATCGTTGTGAGAGTAAAGTTGGCGCAGTGT 6669
Db 4410 AAAAGCGTTAGCTCCTTTCGGTCTCTCCGATCGTTGTGAGAGTAAAGTTGGCGCAGTGT 4469
QY 6670 ATCACTCATGTTATGGCAGCATGCTCAATAATCTCTTACTGTATGCGCATCCGTTAAGATG 6729
Db 4470 ATCACTCATGTTATGGCAGCATGCTCAATAATCTCTTACTGTATGCGCATCCGTTAAGATG 4529

QY	6730	CTTTTCTGTGACTCGGTGAGTACTCAACCAAGTCAATCTGTGAGATAGTGTATGCGGCGAC	6789
Db	4530	CTTTTCTGTGACTCGGTGAGTACTCAACCAAGTCAATCTGTGAGATAGTGTATGCGGCGAC	4589
QY	6790	GAGTTGTCTTTGCCCGCGTCAATACCGGATAATACCGGCCACATAGCAGAACTTTTAA	6849
Db	4590	GAGTTGTCTTTGCCCGCGTCAATACCGGATAATACCGGCCACATAGCAGAACTTTTAA	4649
QY	6850	AGTGCTCATATTGGAATAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT	6909
Db	4650	AGTGCTCATATTGGAATAAGTCTTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT	4709
QY	6910	GAGATCAGTTTCGATGTAAACCCACTCGTGACCCAACTGATCTTCGAGTCTTTTACTTT	6969
Db	4710	GAGATCAGTTTCGATGTAAACCCACTCGTGACCCAACTGATCTTTTACTTTTACTTT	4769
QY	6970	CACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAG	7029
Db	4770	CACCAGCGTTTCTCGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAG	4829
QY	7030	GGCGACACGGAAATGTGTGAATFACTATCACTCTCTCTTTTCAATATTATTGAAAGCATTTA	7089
Db	4830	GGCGACACGGAAATGTGTGAATFACTATCACTCTCTCTTTTCAATATTATTGAAAGCATTTA	4889
QY	7090	TCAGGGTTATTGTCTCATAGCGGATACATATTTGAAATGTATTAGAAAAATAAACAAAT	7149
Db	4890	TCAGGGTTATTGTCTCATAGCGGATACATATTTGAAATGTATTAGAAAAATAAACAAAT	4949
QY	7150	AGGGTTTCCGCGACATTTCCCGAAAGTGCCACCTGACGTCTTAAGAAACCATTTAT	7209
Db	4950	AGGGTTTCCGCGACATTTCCCGAAAGTGCCACCTGACGTCTTAAGAAACCATTTAT	5009
QY	7210	CATGACATTAACTTATAAAATAGCGGTATCAACGAGGCCCTTTCGTCTCGCGGGTTTCGG	7269
Db	5010	CATGACATTAACTTATAAAATAGCGGTATCAACGAGGCCCTTTCGTCTCGCGGGTTTCGG	5069
QY	7270	TGATGACGGTGAATACTCTTGACACATGCAAGTCTCCGAGACGGTCAACGCTGTCTGTA	7329
Db	5070	TGATGACGGTGAATACTCTTGACACATGCAAGTCTCCGAGACGGTCAACGCTGTCTGTA	5129
QY	7330	AGCGGATGCCGGAGCAGACAAGCCCGTCAAGGCGGTCTGACCGGGTGTTCGGCGGGTCTCG	7389
Db	5130	AGCGGATGCCGGAGCAGACAAGCCCGTCAAGGCGGTCTGACCGGGTGTTCGGCGGGTCTCG	5189
QY	7390	GGGCTGGCTTAACTTATCGGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTG	7449
Db	5190	GGGCTGGCTTAACTTATCGGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTG	5249
QY	7450	TGAATAACCGACAGATGCGGTAAAGGAAAAATACCGCATCAGG	7492
Db	5250	TGAATAACCGACAGATGCGGTAAAGGAAAAATACCGCATCAGG	5292

RESULT 4

```

US-08-793-610-2
; Sequence 2, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARRERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaïdo, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C..
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

Db 2133 GTCCCGGGTACCCGGTTCCTCAATAAAACCCCTCTTGCACTGTCATCCGACTCGTGGTCT 2192
Qy 4449 CGCTGTTCTTGGAGGGTCT-CTCTGAGTGATTGACTACCGTCAAGC-GGGGTCTTTTCA 4506
Db 2193 CGCTGTTCTTGGAGGGTCTCTCTGAGTGATTGACTACCGGTCAAGCGGGGTCTTTTCA 2252
Qy 4507 GTTTCCTCCCACTTACACAGGTCTCACTAAACATTCCTGATGTGCGGAGGAGCTCCGTGAG 4566
Db 2253 GTTTCCTCCCACTTACACAGGTCTCACTAAACATTCCTGATGTGCGGAGGAGCTCCGTGAG 2312
Qy 4567 CCGGTTTTTGTATTATAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 4626
Db 2313 CCGGTTTTTGTATTATAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 2372
Qy 4627 TGAATCTCGGCTTTATAAAGATGTTGAAGGCTCTGTGGACTATCTGCCACAGACTT 4686
Db 2373 TGAATCTCGGCTTTATAAAGATGTTGAAGGCTCTGTGGACTATCTGCCACAGACT-T 2431
Qy 4687 TTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGC 4746
Db 2432 TTAAAGATTTTATGCTCTCTGGATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGC 2491
Qy 4747 TGCTTCTCCGATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCACGCGATTGC 4806
Db 2492 TGCTTCTCCGATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCACGCGATTGC 2551
Qy 4807 AGAGTAGTTAACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 4866
Db 2552 AGAGTAGTTAACTCTGAGGCGAGGCTTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 2611
Qy 4867 ATTTTCCCAATAAAAGGGTCTTGCCCTGTGTACAGACAGATCGGAATCTAGATGTCAT 4926
Db 2612 ATTTTCCCAATAAAAGGGTCTTGCCCTGTGTACAGACAGATCGGAATCTAGATGTCAT 2671
Qy 4927 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCGCTAGAGATCCCTC 4986
Db 2672 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCGCTAGAGATCCCTC 2731
Qy 4987 TAGCCGAGAGCATCTGCGCGGTACCGAGCTCGAATTCGTAATCATGGTCTATAGTGT 5046
Db 2732 TAGCCGAGAGCATCTGCGCGGTACCGAGCTCGAATTCGTAATCATGGTCTATAGTGT 2791
Qy 5047 TTC-----CTGTGTGAATTTGTATTCGCTCACAATTCACACACATACGA 5093
Db 2792 ATCCGCGCATCACCGATGGGAGATCGGGTCCGCACTTCGGGCTCATGAGCGCTGT 2851
Qy 5094 GCCGGAAGCATAAAGTGTAAAGCTTGGGGTGCTTAATGAGT----- 5134
Db 2852 TTCGGGTGGTATGTGTCAGGCCCCGGTGGCGGGGACTGTTGGGCGCCATCTCCTTG 2911
Qy 5135 -----GAGCTAACTCACATTAATTTGGTTGGCTCACTGCC 5170
Db 2912 CATGCAACCATTCCTTGGCGGGCGGTGCTCAACGGGCTCAACCTACTAGGGGCTGCTTC 2971
Qy 5171 CGCTTTCAGTCCGGAACCTGTGTCAGCTGCAATTAATGAATCGGCAACCGCGGG 5230
Db 2972 CTAATCAGAGTCCGATTAAGGAGAGCGTCTGCAATTAATGAATCGGCAACCGCGGG 3031
Qy 5231 GAGAGCGGTTTTCGTTATTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTCGGCTC 5290
Db 3032 GAGAGCGGTTTTCGTTATTTGGGCGCTTTCCGCTTCTCGCTCACTGACTCGCTCGGCTC 3091
Qy 5291 GGTCTGTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGGTAATACGGTTATCCAC 5350
Db 3092 GGTCTGTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAAGGCGGTAATACGGTTATCCAC 3151
Qy 5351 AGAATCAGGGGATAACGAGGAAGAAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAA 5410
Db 3152 AGAATCAGGGGATAACGAGGAAGAAACATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAA 3211
Qy 5411 CCGTAAAAAGGCGCGTCTGCGTTTTTCCATAGGCTCCGCCCCCTCGACGAGCATCA 5470
Db 3212 CCGTAAAAAGGCGCGTCTGCGTTTTTCCATAGGCTCCGCCCCCTCGACGAGCATCA 3271

Qy 5471 CAAAAATCGACGCTCAAGTCAAGGTGGGAAACCCGACAGACTATATAAGATACACGGC 5530
Db 3272 CAAAAATCGACGCTCAAGTCAAGGTGGGAAACCCGACAGACTATATAAGATACACGGC 3331
Qy 5531 GTTTCCTCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCGAGCCCTGCGCTTACCGGATA 5590
Db 3332 GTTTCCTCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCGAGCCCTGCGCTTACCGGATA 3391
Qy 5591 CTTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGTCAACGCTGTAGGTA 5650
Db 3392 CTTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTA 3451
Qy 5651 TCTCAGTTTCGGTGTAGTTCGCTCCAAAGCTGGGCTGTGTGACGACACCCCGCTTCA 5710
Db 3452 TCTCAGTTTCGGTGTAGTTCGCTCCAAAGCTGGGCTGTGTGACGACACCCCGCTTCA 3511
Qy 5711 GCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAGACACA 5770
Db 3512 GCCGACCGCTGCGCTTTATCCGGTAACTATCGTCTTGAGTCCAAACCCGGTAAGACACA 3571
Qy 5771 CTTATCGGCACTGGCAGACGCCACTGTGTAAACAGGATTTAGCAGAGGATGTAGGCGG 5830
Db 3572 CTTATCGGCACTGGCAGACGCCACTGTGTAAACAGGATTTAGCAGAGGATGTAGGCGG 3631
Qy 5831 TGCTACAGAGTTCTTGAAGTGGTGCCTAACTACCGCTACACTAGAACGACAGTATTTGG 5890
Db 3632 TGCTACAGAGTTCTTGAAGTGGTGCCTAACTACCGCTACACTAGAACGACAGTATTTGG 3691
Qy 5891 TATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGAAAAAAGAGTTGGTAGCTCTTGATCCCG 5950
Db 3692 TATCTGCGCTCTGCTGAAAGCCAGTTTACCTTCGAAAAAAGAGTTGGTAGCTCTTGATCCCG 3751
Qy 5951 CAAAACAAACCCGCTGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGATTAACGCGAG 6010
Db 3752 CAAAACAAACCCGCTGTGTAGCGGTGTTTTTTTGTTCGAAAGCAGAGATTAACGCGAG 3811
Qy 6011 AAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAA 6070
Db 3812 AAAAAAGGATCTCAAGAGATCCCTTTGATCTTTTCTACGGGCTGACGCTCAGTGAA 3871
Qy 6071 CGAAAACTCAGCTTAAAGGATTTTGTCTCATGAGATTTATCAAAAAGGATCTTCAACCTAGAT 6130
Db 3872 CGAAAACTCAGCTTAAAGGATTTTGTCTCATGAGATTTATCAAAAAGGATCTTCAACCTAGAT 3931
Qy 6131 CTTTTTAAATTTAAATGAAATTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTC 6190
Db 3932 CTTTTTAAATTTAAATGAAATTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTC 3991
Qy 6191 TGACAGTTTACCAATGCTTAATCAGTGAGGCACTTCTCAGCGATCTGCTATTTTCGTTTC 6250
Db 3992 TGACAGTTTACCAATGCTTAATCAGTGAGGCACTTCTCAGCGATCTGCTATTTTCGTTTC 4051
Qy 6251 ATCCATAGTTGCTGACTCCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 6310
Db 4052 ATCCATAGTTGCTGACTCCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 4111
Qy 6311 TGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGCTCCAGATTTATFAGC 6370
Db 4112 TGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGCTCCAGATTTATFAGC 4171
Qy 6371 AATTAACCAAGCCGAGGAGGCGGAGAGTGGTCTGCAACTTTATCCGCTTC 6430
Db 4172 AATTAACCAAGCCGAGGAGGCGGAGAGTGGTCTGCAACTTTATCCGCTTC 4231
Qy 6431 CATCCAGTCTATTAAATTTGTTCCCGGAAAGCTAGAGTAAAGTAGTTCCGCACTTAATAGTTT 6490
Db 4232 CATCCAGTCTATTAAATTTGTTCCCGGAAAGCTAGAGTAAAGTAGTTCCGCACTTAATAGTTT 4291
Qy 6491 GCGCAACGTTGTTGCCATTTGCTACAGGC-TGCTGTGTACGCTCGTCTGTTGGTATGCG 6549
Db 4292 GCGCAACGTTGTTGCCATTTGCTACAGGCATCGTGTGTACGCTCGTCTGTTGGTATGCG 4351

```
QY 6550 TTCATTGAGCTCCGGTCCCAACGATCAAGCGAGTTACATGATCCCCATGTTGTGCAA 6609
Db 4352 TTCAATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACATGATCCCCATGTTGTGCAA 4411
QY 6610 AAAAGCGGTTAGCTCTTCCTCGCTCCCGATCGTTGTGAGAAAGTAAGTTGGCGCGAGTGT 6669
Db 4412 AAAAGCGGTTAGCTCTTCCTCGCTCCCGATCGTTGTGAGAAAGTAAGTTGGCGCGAGTGT 4471
QY 6670 ATCACTCATGTTATGCGAGCACTGCAATAATCTCTTACTGTCTCATGCGCATCCGTAAGATG 6729
Db 4472 ATCACTCATGTTATGCGAGCACTGCAATAATCTCTTACTGTCTCATGCGCATCCGTAAGATG 4531
QY 6730 CTTTTCTGTGAGTGGTACTCAACCAAGTCAATCTGAGAAAGTATGATGCGCGGAC 6789
Db 4532 CTTTTCTGTGAGTGGTACTCAACCAAGTCAATCTGAGAAAGTATGATGCGCGGAC 4591
QY 6790 GAGTGTCTTTCGCGCGGTCAATACGGGATAATACCGCGCACATAGCAGAACTTTAAA 6849
Db 4592 GAGTGTCTTTCGCGCGGTCAATACGGGATAATACCGCGCACATAGCAGAACTTTAAA 4651
QY 6850 AGTGCTCATCATGGAACACGTTCTTCGGGCGGAAACCTCTCAAGGATCTTTACCGCTGTT 6909
Db 4652 AGTGCTCATCATGGAACACGTTCTTCGGGCGGAAACCTCTCAAGGATCTTTACCGCTGTT 4711
QY 6910 GAGATCCAGTTCGATGTAACCACTCGTGCACCCACTGATCTTCAGCATCTTTACTTT 6969
Db 4712 GAGATCCAGTTCGATGTAACCACTCGTGCACCCACTGATCTTCAGCATCTTTACTTT 4771
QY 6970 CACGAGCGTTCTGGGTGAGCAAAACAGGAAGCAAAATCCGCAAAAGGGAATAAG 7029
Db 4772 CACGAGCGTTCTGGGTGAGCAAAACAGGAAGCAAAATCCGCAAAAGGGAATAAG 4831
QY 7030 GCGGACACGGAATGTTGAATACCTACTCTCTCTTTTCAATATATTGAAAGCATTTA 7089
Db 4832 GCGGACACGGAATGTTGAATACCTACTCTCTCTTTTCAATATATTGAAAGCATTTA 4891
QY 7090 TCAGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAACAAT 7149
Db 4892 TCAGGTTATTGTTCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATAACAAT 4951
QY 7150 AGGGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGCTTAAGAAACCATTTAT 7209
Db 4952 AGGGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGCTTAAGAAACCATTTAT 5011
QY 7210 CATGACATTAACCTATAAAATAGCGGTATCACGAGGCGCTTTGCTCTCGCGCGTTTCGG 7269
Db 5012 CATGACATTAACCTATAAAATAGCGGTATCACGAGGCGCTTTGCTCTCGCGCGTTTCGG 5071
QY 7270 TGATGACGTTGAAACCTCTGACACATGCACTCCCGGAGCGGTCAAGCTTTGCTGTA 7329
Db 5072 TGATGACGTTGAAACCTCTGACACATGCACTCCCGGAGCGGTCAAGCTTTGCTGTA 5131
QY 7330 AGCGGATCCGCGGAGCAGACAGCCGTCAGGCGCGCTCAGCGGTTGTCGGGTCG 7389
Db 5132 AGCGGATCCGCGGAGCAGACAGCCGTCAGGCGCGCTCAGCGGTTGTCGGGTCG 5191
QY 7390 GGGCTGGCTTAACCTATGCGGATCAGAGCAGATTGTTACTGAGAGTGCACCATATGCGGTG 7449
Db 5192 GGGCTGGCTTAACCTATGCGGATCAGAGCAGATTGTTACTGAGAGTGCACCATATGCGGTG 5251
QY 7450 TGAATACCGCACAGATCGTAAGGAGAAAAATACCGCATCAGG 7492
Db 5252 TGAATACCGCACAGATCGTAAGGAGAAAAATACCGCATCAGG 5294
```

RESULT 5

```
US-08-793-610-1
; Sequence 1, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
```

```
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: F1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-1
```

```
Query Match 40.3%; Score 3083.4; DB 2; Length 5323;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 3400; Conservative 0; Mismatches 191; Indels 112; Gaps 9;

QY 3868 ACAGGACATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAG 3927
Db 1655 ACAGGACATAGATAAAATAAAAGATTTTATTAGTCTCCAGAAAAAGGGGGAATGAAG 1714
QY 3928 ACCCCACCTGTAGGTTTGGCAAGC-----TAGCTTAAAGTAAGCCCAT 3969
Db 1715 ACCCCACCTGTAGGTTTGGCAAGCTAGAGTCGCTTAGCTGATAGCCGATTAAGCCAT 1774
QY 3970 TTTCAGAGGATGCG-AAAATACATACTAGAGATAGAGAAAGTTCCAGATCAAGGTTAGGAA 4028
Db 1775 TTTCAGAGGATGCGAAAAATACCAACCAAGATAGGAAAGTTCCAGATCAAGGCGGTA 1834
QY 4029 CAGAGACAGCAGAAATATGGGCCAAAACAGATATCTGTGTAGAGATTCCTGCTGCCCGC 4089
Db 1835 CATGAAAAATAGCTAACGTTGGGCCAAAACAGATATCTGCGGTGAGCAGTTTCGCGCCCGG 1894
QY 4089 TCAGGGCCAAACAGTTTGGAAACAGGAGATATGGGCCAAACAGATATCTGTGTAGAGC 4148
Db 1895 CCGGGCCAAACAGATG-----TCACCGC 1921
QY 4149 AGTTCTCTGCCCGCTCAGGGCCCAAGAACAGATGTCCTCCAGATCGCGTCCCGCCCTCAG 4208
Db 1922 AGTTCTGCCCGCTCCGCGCCCAAGAACAGATGTCCTCCAGATATGCGCCCAACCTCAG 1981
QY 4209 CAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGT 4268
Db 1981 CAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAATGACCTGT 4268
```

Db 1982 CAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAAATGACCCCTGC 2041
Qy 4269 GCCTTATTTGAATTAACCAATCAGTTCGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTGCC 4328
Db 2042 GCCTTATTTGAATTAACCAATCAGCTGCTTCTCGCTTCTGTTGCGGCGCTTCTGCTGCC 2101
Qy 4329 CGAGCTCAATAAAGAGGCCCAACACCCCTCACTCGGCGCGCCAGTCTCTCCGATAGACTGC 4388
Db 2102 CGAGCTCATAAAGAGCTCAACACCCCTCACTCGGCGCGCCAGTCTCTCCGATGACTGA 2161
Qy 4389 GTGCGCGGGTACCCGCTGTTCTCAATAAACCCTCTTGCAAGTTGCAATCCGACTCGTGGTCT 4448
Db 2162 GTGCGCGGGTACCCGCTGTTCTCAATAAACCCTCTTGCAAGTTGCAATCCGACTCGTGGTCT 2221
Qy 4449 CGCTGTTCTCGGAGGGTCT-CTCTGAGTGATGACTACCCGTCAGC-GGGGCTTTTCA 4506
Db 2222 CGCTGTTCTCGGAGGGTCTCTCTGAGTGATGACTACCCGTCAGGCGGGGTCTTTCA 2281
Qy 4507 GTTTCCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGAGCTCCGTCAG 4566
Db 2282 GTTTCCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGAGCTCCGTCAG 2341
Qy 4567 CCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 4626
Db 2342 CCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTACATCC 2401
Qy 4627 TGACTCTCGGCTTATAAAGATGTTGAAGGCTCTGGAAGTATCTGCACTATCTGCCACAGCTT 4686
Db 2402 TGACTCTCGGCTTATAAAGATGTTGAAGGCTCTGGAAGTATCTGCACTATCTGCCACAGC-T 2460
Qy 4687 TTTAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGCTATTTTGC 4746
Db 2461 TTTAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGCTATTTTGC 2520
Qy 4747 TGCTTCTCCGATTTAATTTCTAGTTTGCATCTCCCTTCTGAGAGCACGCGATTCG 4806
Db 2521 TGCTTCTCCGATTTAATTTCTAGTTTGCATCTCCCTTCTGAGAGCACGCGATTCG 2580
Qy 4807 AGAGTAGTTTAACTCTGAGGCGAGCTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 4866
Db 2581 AGAGTAGTTTAACTCTGAGGCGAGCTCTGTGAAGAGTTGCTGGGCTCAGTGTGAG 2640
Qy 4867 ATTTTGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGAATCTAGAGTGCAT 4926
Db 2641 ATTTTGCCATAAAGGGGCTCTGCCCTGTGTACAGACAGATCGAATCTAGAGTGCAT 2700
Qy 4927 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCTCAGAGATCTCTC 4986
Db 2701 ACTCAGAGTCCCGGGTTCGCGGCTCTGATCTCAGGGCATCTTTGCTCAGAGATCTCTC 2760
Qy 4987 TAGCGGAGCGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCTCATAGCTGT 5046
Db 2761 TAGCGGAGCGATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCTCATAGCTGT 2820
Qy 5047 TTTCTCTGTTGAAATCTGTTATCGCTCACAATTTCCACACAACATACGA 5093
Db 2821 ATTCGCGACATCACCGATGGGGAAGATCGGGTTCGCCATTCGCGGCTCATGAGCGCTGT 2880
Qy 5094 GCGGGAAGCATAAAGTGTAAAGCTGCGGGTGCCTAATGAGT----- 5134
Db 2881 TTTGCGGTGGTATGGTGGCAGGCCCGGCGCGGGGACTGTGGGCGCCATCTCCTTTG 2940
Qy 5135 -----GAGCTAACTCAATTAATTTGCGTTCGCGTCACTGCC 5170
Db 2941 CATGCACCATTCCTTTCGCGCGGGTGTCTCAACGGGCTCAACCTACTACTGCGGCTCTTC 3000
Qy 5171 CGCTTTTCAGTCCGGGAACCTGTGTCGCGAGTGCATTAATGAATCGGCAACGCGCGGG 5230
Db 3001 CTAAATCAGAGTCCGATTAAGGAGAGCGCTCTGCAATTAATGAATCGGCAACGCGCGGG 3060
Qy 5231 GAGAGCGGTTTGGGATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 5290
Db 3061 GAGAGCGGTTTGGGATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTC 3120

Qy 5291 GGTGTTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTTATCCAC 5350
Db 3121 GGTGTTTCGGCTCGGCGAGCGGTATCAGCTCACTCAAAGCGGTAAATACCGTTATCCAC 3180
Qy 5351 AGAATCAGGGGATTAACGCAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAA 5410
Db 3181 AGAATCAGGGGATTAACGCAAGAAACATGTGAGCAAAAGCCAGCAAAAGCCAGGAA 3240
Qy 5411 CCGTAAAGGCGCGGTTGCTGGGTTTTTCCATAGGCTCGCCCCCTGAGCAGCATCA 5470
Db 3241 CCGTAAAGGCGCGGTTGCTGGGTTTTTCCATAGGCTCGCCCCCTGAGCAGCATCA 3300
Qy 5471 CAAAAATCGACGCTCAAGTCAGAGTGCGGAAACCCGACAGACTATATAAGATACACGCG 5530
Db 3301 CAAAAATCGACGCTCAAGTCAGAGTGCGGAAACCCGACAGACTATATAAGATACACGCG 3360
Qy 5531 GTTTCCTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCGCTTACCGGATA 5590
Db 3361 GTTTCCTCCCTGGAAAGCTCCCTCGTGCCTCTCTGTTCCGACCCCTGCGCTTACCGGATA 3420
Qy 5591 CCGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGTAGTGA 5650
Db 3421 CCGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGTAGTGA 3480
Qy 5651 TCTCAGTTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGACAGAAACCCCGCTTCA 5710
Db 3481 TCTCAGTTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGACAGAAACCCCGCTTCA 3540
Qy 5711 GCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCCGTAAAGACAGA 5770
Db 3541 GCCGACCGCTGCGCTTATCCGGTAACTATCGTCTGAGTCCAAACCCCGTAAAGACAGA 3600
Qy 5771 CTTATCGCCTGCGCAGCAGCCTGCTGTAACAGGATAGCAGCGAGGTATGTAGGCGG 5830
Db 3601 CTTATCGCCTGCGCAGCAGCCTGCTGTAACAGGATAGCAGCGAGGTATGTAGGCGG 3660
Qy 5831 TGCTACAGATTTCTGAAGTGTGGCTTAACTTACGGCTACACTAGAAGACAGATTTGG 5890
Db 3661 TGCTACAGATTTCTGAAGTGTGGCTTAACTTACGGCTACACTAGAAGACAGATTTGG 3720
Qy 5891 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGGATGCTGCTGTGATCCGG 5950
Db 3721 TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGGATGCTGCTGTGATCCGG 3780
Qy 5951 CAAACAAACCAACCGCTGCTGAGCGGTGTTTTTGTGTTGCAAGCAGCAGATTAACGCGCAG 6010
Db 3781 CAAACAAACCAACCGCTGCTGAGCGGTGTTTTTGTGTTGCAAGCAGCAGATTAACGCGCAG 3840
Qy 6011 AAAAAAGGATCTCAAGAGATCTTTTTCATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 6070
Db 3841 AAAAAAGGATCTCAAGAGATCTTTTTCATCTTTTCTACGGGCTCTGACGCTCAGTGGAA 3900
Qy 6071 CGAAAACTCAGCTTAAGGGATTTTGGTTCATGAGATTAATAAAGGATCTTCCACCTAGAT 6130
Db 3901 CGAAAACTCAGCTTAAGGGATTTTGGTTCATGAGATTAATAAAGGATCTTCCACCTAGAT 3960
Qy 6131 CTTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATGATGATAAAGTGTGTC 6190
Db 3961 CTTTTTAAATTAATAAGATTTTAAATCAATCTAAAGTATATATGATGATAAAGTGTGTC 4020
Qy 6191 TGACAGTTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTGTCTATTTTCGTTTC 6250
Db 4021 TGACAGTTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTGTCTATTTTCGTTTC 4080
Qy 6251 ATCCCATAGTGTGCTGACTCCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 6310
Db 4081 ATCCCATAGTGTGCTGACTCCCCCTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC 4140
Qy 6311 TGCGCCAGTGTGCAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTTATCAGC 6370
Db 4141 TGCGCCAGTGTGCAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTTATCAGC 4200

498 TTTGAACTAACCAATCAGTTTCGCTTCTCGCTTCTGTTTCGGGGCTTCCGCTCTCCGAGCT 557
 965 CAATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCCCTCCGATAGACTCGGTCGCC 1024
 558 CAATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCCCTCCGATAGACTCGGTCGCC 617
 1025 CGGTTACCCGTTATTCCTCAATAAAGCCTCTTGTGTTTGTGATCCGAATCGTGGACTCGCTG 1084
 618 CGGTTACCCGTTATTCCTCAATAAAGCCTCTTGTGTTTGTGATCCGAATCGTGGACTCGCTG 677
 1085 ATCTTGGGAGGCTTCCTCAGATTGATTGACTGCCCCAATCGTGGGGTCTTCAATTGGA 1144
 678 TTTCTTGGGAGGCTTCCTCAGATTGATTGACTGCCCCAATCGTGGGGTCTTCAATTGGA 737
 1145 GGTTCACCCAGATTGGAGACCCCTCGCCAGGGACCCAGCCGCCCGCCGGGAGGTA 1204
 738 GGTTCGTCGGGATTGGAGACCCCTCGCCAGGGACCCAGCCGCCCGCCGGGAGGTA 796
 1205 AGCTGGCCAGCAACTTATCTGTGCTGTCTCGGATTGCTAGTGTCTATGACTGATTTTATG 1264
 797 AGCTGGCCAGCAACTTATCTGTGCTGTCTCGGATTGCTAGTGTCTATGATTGATTATG 856
 1265 CGCTGCTGCTGCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGT 1324
 857 CGCTGCTGCTGCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGT 916
 1325 GAGGAGTTCGGACACCCCGCCGACCCCTCGGAGACGTCCTCGGAGACTTCGGGGCCGT 1384
 917 GAGGAGTTCGGACACCCCGCCGACCCCTCGGAGACGTCCTCGGAGACTTCGGGGCCGT 976
 1385 TTTTGTGGCCGACCTGAGTCTTAAATCCCGATCGTTTGTAGACTCTTTTGTGACCCGCC 1444
 977 TTTTGTGGCCGACCTGAGGAGGAGTCTGATGTGAATCCGACCC----- 1022
 1445 CTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTCCCGCTCC 1504
 1023 -----CGTCAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTCCCGCTCC 1077
 1505 GTCTGAATTTTGTCTTTCGGTTTGGACCGAAGCCGCGCGCTGTGT---CTGCTG 1561
 1078 GTCTGAATTTTGTCTTTCGGTTTGGAACCGAAGCCGCGCGCTGTGTCTGCTCAGCGCTG 1137
 1562 CAGCATCGTCTGTGTGTCTGTCTGACTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1621
 1138 CAGCATCGTCTGTGTGTCTGTCTGACTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1197
 1622 CCCGGCTAGACTGTTTACCACTCCCTTAAAGTTTGTGACTTGTAGTCTACTGGAAAGATGCGA 1681
 1198 CC-----AGACTGTTTACCACTCCCTTAAAGTTTGTGACTTGTAGTCTACTGGAAAGATGCGA 1251
 1682 GGGATCGCTCACAACCACTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGTCTC 1741
 1252 GGGATCGCTCACAACCACTCGGTAGATGTCAAGAGAGACGTTGGGTTACCTTCTGTCTC 1311
 1742 TGCAGATCGCCAACTTAAAGTTCGGATGGCGCGGAGACGCACTTAAACCGAGACCT 1801
 1312 TGCAGATGGCCAACTTAAAGTTCGGATGGCGCGGAGACGCACTTAAACCGAGACCT 1371
 1802 CATCACCCAGGTTAAGATCAAGGTTCTTTTCACTGGCCCGCATGGACACCCAGACAGGT 1861
 1372 CATCACCCAGGTTAAGATCAAGGTTCTTTTCACTGGCCCGCATGGACACCCAGACAGGT 1431
 1862 CCCCTACATCGTGAACCTGGAGAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGCCCTT 1921
 1432 CCCCTACATCGTGAACCTGGAGAGCTTGGCTTTTGAACCCCTTCCCTGGGTCAAGCCCTT 1491
 1922 TGFACACCTTAAGCTTCCGCTCTCTTCTCCATCCGCGCCGCTCTCTCCCTTGAACC 1981
 1492 TGFACACCTTAAGCTTCCGCTCTCTTCTCCATCCGCGCCGCTCTCTCCCTTGAACC 1551
 1982 TCCTCGTTCGACCCCGCTCGATCTCTCTCTTATCCAGCCCTCACTCTCTCTCTAGGGCG 2041

1552 TCCTCGTTCGACCCCGCTCGATCTCTCCCTTTATCCAGCCCTCACTCTCTCTCTAGGGCG 1611
 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTCA 2101
 1612 C-----GGAAATTCGGATCTGA 1627
 2102 CCTCGACATGACAAGAGTTACTTAACAGCCCTCTCTCCAAGCTCACTTACAGGCTTCTAC 2161
 1628 TCAAGAGACAGGATGAGGATGTTTTCGATGATTGAACAAGATGATTGACGAGGTTTC 1687
 2162 TTAGTTCAGACAGAAAGTCTGGAGACTCTGGGGGAGCTTCAACAAGAACTTACAGACGA 2221
 1688 TCCGCGCGCTTG-----GTGGAGAGGCTATTTCGGTATGACTTGGGCACAA 1733
 2222 CGGTTGTACCTCACCTTACCGAGTCGGGACACAGTGTGGTCCGCGACACAGACT 2281
 1734 CAGACAATCGGCTGTCTGATGCCCGCTGTTCCGGCTGTGAGCGAGGGGCGCCGTT 1793
 2282 AAGAACTTAGAACCTCGCTGGAAGAGACTTACACAGTCTCTGCTGACCAACCCCGCCG 2341
 1794 CTTTGTGTCAAGACCGACTCTGTCGGTGCCTGTAATGAATGC----- 1836
 2342 CTCAAAGTAGAGCGCATCGAGCTTGGATACACGCCGCCACAGTGAAGGCTGCCGACCCC 2401
 1837 -----AGGACAGGCGAGCGCGCTATCGTGGCTGCGCACGACGCGCTTCTCTTGGCGAC 1891
 2402 GGGGTGACCATCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTCTTGGTAG 2461
 1892 TGTGCTGACAGTGTCTGTAAGCGGAGGAGTGTGGCTGTATTTGGGCGAAGTCCCGG 1951
 2462 CAACAGCTTACAGTGTCTCACTCCAGCTCAGCTCAGCGACGACGACGACGCTGAGCG 2521
 1952 GCAGATCTCTGTGTCATCTCACTTGTCTCTGCGGAGAAAGTATCCATCATGCTGTATG 2011
 2522 CAGCGTGGGTGACAGAGTACCATCACTGTAAAGCCAGTCAAGATGTGGGTACTTCTG 2581
 2012 AATCGCGGCTGCTAGCTTGTATCGGCTACCTGCGCATTCGACCAACCAAGCGAACA 2071
 2582 TAGCTTGTGTAACAGAGACGCGAGTAAAGCTTCCAAAGCTCTGATCTACTTGGACATCA 2641
 2072 TCGCATCAGCGAGACGCTACTCGGATGGAAGCCGCTTGTGTCGATCAGGATGATCTGA 2131
 2642 CCGCGCACACTGCTGTGCAAGCAGATTTCAGCGGTAGCGGTAGCGGTACCGACTTCACCT 2701
 2132 CGAAGACATCAGGGCTCGCGCCAGCCGAACTGT-----TCGCCAGC 2175
 2702 TCACCATCAGCAGCTCCTCAGCAGAGGACATCGCCACCTACTACTTGCAGCAATATAGCC 2761
 2176 TCAAGCGCGATGCGCGAGGAGATCTGCTGTGACCCCATGGCGATGCTGCTTGC 2235
 2762 TCTATCGGCTGTTCCGCCAAGGGACCAAGTGGAAATCAACAGAGTGGCTCAGGATCGG 2821
 2236 CGAATATCATGGTGGAAATGGCCGCTTTTCTGGATTCTGAGCTGTGGCGGCTGGGTG 2295
 2822 GTGATCGGCTCTGCTGCTCAGGATCGGAGTCAACTGGTGGAGAGCGGTGGAGGTG 2881
 2296 TGGCGGACCGCTATCAGAGACATAGCTTGGCTACCCGCTGATATGCTGAGAGCTTGGCG 2355
 2882 TTGTGCAACCTGGCGCTCCCTGCGCTCTCTGCTCCGCACTCTGGCTTCGATTTCAACA 2941
 2356 GCGAATGGCTGACCGCTTCTCTGCTTACCGTATCGCCCTCC----- 2402
 2942 CATATTGATGATTGGGTGAGACAGGACCTGGAAAGGTTCTGAGTGGATTGGAGAA 3001
 2403 -----GATTCGACGCGCATCGCTTCTCTATCGCTTCTTTCACGAG 2441
 3002 TTATCCAGATAGTAGTACGATTAATGCTGCGCTCTCTAAAGGATAGATTTACAATAT 3061
 2442 TTCTTCTGAGCGGACTCTGGGTTGGAATGACCGACCAAGCGACGCCCACTGCCAT 2501
 3062 CGCGAGACCAAGCAACATTTGTTCTTCTGCAATGGACGCTGAGACCCGAGACA 3121
 2502 CACGAGATTTGATTTCCACCGCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCC 2561

QY 3122 CCGGGCTTAATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGGTTTGTGTTATTTGGGGCC 3181
DB |||||
3182 AAGGACCCCGGTCAACGCTCTCCAGTGTAAAGCCACACGACGCGCGGCGGAC 3241
DB |||||
3242 CAACACCGCGGCCCAATCGCGTCCGAGCCCTGTCTCCGCGCCAGAGCGGCTCGGC 3301
DB |||||
3302 CAGCGCGGGGGCGCAGTGCACAGAGGGGCTGGACTTCGCCCTCGATCCCAACTCT 3361
DB |||||
3362 GCTACCTCTGATGGAATCTCTTCACTATGATGTCATCTCACTGCTTGTCTCTGA 3421
DB |||||
3422 GAGTGAAGTTTCAAGCAGGCGCAGAGCCCGCGTACCGAGCGGCGCAGAACAGCTCT 3481
DB |||||
3482 ATAACGAGCTCAATCTAGGACGAAGAGAGAGTACGATGTTTGTGCAAGAGACGTGCC 3541
DB |||||
3542 GGGACCTGAGATGGGGGGAAGCGAGAGAGGAGAACCTTCAGAGAGCGCTGTACAATG 3601
DB |||||
3602 AACTGCAGAAAGATAAGATGGCGAGGCGCTACAGTGAGATGGGATGAAGCGGAGCGCC 3661
DB |||||
3662 GGAGGGCAAGGGGCAAGATGGCTTTTACAGGGTCTCAGTACAGCCACCAAGGACCT 3721
DB |||||
3722 ACAGCGCCCTTACATGACGAGCGCCCTGCCCCCTCGCTAATCTCGACGCGCGCGGATCCGG 3781
DB |||||
3782 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCAGTGTGACTCAACA 3841
DB |||||
3842 ATATCACCAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTAG 3901
DB |||||
3902 TCTCAGAAAAGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGT 3961
DB |||||
3962 AACGCCATTTTGAAGGATCG-AAAATACATAACTGAGATAGAGAGTTTCAAGTCAAG 4020
DB |||||
4021 GTTAGGAACAGA-GAGACAGCAGATATGGGCCAAACAGGATATCTGTGTGAAGCAGTTC 4079
DB |||||
4080 CTGCCCC-GCTCAGGGCCAAAGACGTTGGAAACAGGAGATATGGGCCAAACAGGATATC 4138
DB |||||
4139 TGTGTAGCAGTTCTCTCCCGCTCAGGCGCAGACAGATGTTCCCGAGTCCGCTC 4198
DB |||||
4200 TGTGTAGCAGTTCTCTCCCGCTCAGGCGCAGACAGATGTTCCCGAGTCCGCTC 4257

QY 4199 CCGCCCTCAGCAGTTTCTTAGAGAACCATCAGATGTTTTCAGGGTGCCTCCCAAGACCTGAA 4258
DB |||||
4259 ATGACCCCTGCTCTTATTTGAACCTAACTAAGTTCGCTTCTCGCTTCTGTTCCGGCCG 4318
DB |||||
4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCGCCAGTCTCTCC 4378
DB |||||
4379 GATAGACTGCGTCCCGGGTACCGGTCTCTCAATAAACCCTCTTTCGAGTTGATCCGA 4438
DB |||||
4439 CTGCTGCTCGCTCTCTTCTTGGGAGGCTCT-CTCTGAGTGTATGACTACCGCTCAGCGG 4497
DB |||||
4498 GGTCTTTCAGTCTTCTCCACCTTACACAGTCTCACTAACTTCTGATGTGCCGAGGA 4557
DB |||||
4558 CTCCGTACGCGGTTTGTATTAATAAATGCAAGACAGTGTCCCTTCAAGCCAG 4617
DB |||||
4618 ACTACATCTGACTCTCGGCTTTATAAAGATGTTGAAGGCTCTGTGGACTATCTGCC 4677
DB |||||
4678 ACAGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCTGT 4737
DB |||||
4738 CTATTTTGTGCTGCTTCTCCGATTTTAAATTTCTAGTTTGCCTCTCCCTTCTGAGAGCAC 4797
DB |||||
4798 GGCATGTCAGAGTAGTTAATACTCTGAGGCGAGCTTCTGTGAAGAGTGTGCTGGGT 4857
DB |||||
4858 CAGTGTGAGATTTTGGCCATAAAGGGGTCTGCGCCCTGTGTGACAGACAGATCGGAATCT 4917
DB |||||
4918 AGAGTCATCTCAGAGTCCCGGTTTCGGGCTCTGATCTCAGGCGATCTTTTGCCTA 4977
DB |||||
4978 GAGATCTCTACGCGGACGATCGTGGCGGGTACCGAGCTCGAATTCGTTAATCATGTT 5037
DB |||||
5038 CATAGCTTGTCTGTGTAATTTGTTATCGCTCACAATTTCCACACAACATACGAGCG 5097
DB |||||
5098 GAAGCATAAAGTGTAAAGCCTTGGGGTGCCTAATGAGTGAGCTAACTCAATTAATTCGT 5157
DB |||||
5158 TGCCTCACTGCCCGCTTTTCCAGTCCGGAACCTGTCTGTCGCGAGCTGATTAATGAATCG 5217
DB |||||
5218 GCCAACCGCGGGGAGAGGGGTTTGGTATTGGCGCTCTTCCGCTTCTCCGCTCCTGCTC 5277
DB |||||
5278 ACTCGCTCGGCTCGTTCGGTTCGGTTCGGCGGAGCGGATATCAGCTCACTCAAAAGCGGTAA 5337


```

; STREET: 624 Ninth Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,312
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: CHANG=112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-935-312-1

Query Match      40.0%; Score 3062.2; DB 3; Length 6145;
Best Local Similarity 71.7%; Pred. No. 0;
Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;

QY 549 TTTTGAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAAACGCCATTTTGCAAG 608
Db 144 TTTTGAAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAGTAAACGCCACTTTGCAAG 201

QY 609 CATGG-AAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTTAGGAACACAGAGAC 667
Db 202 CATGGAAAATACATAACTGGAATAGAAAAGTTTCAGATCAAGGTCAGGAACAAGAAAC 261

QY 668 AGCAGAAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTTCCTGCCCC-GCTCAGGGCC 726
Db 262 AGCTGAATA---CCAAACAGGATATCTGTGGTAAGCGTTCTCTGCCCCCGGCTCAGGGCC 317

QY 727 AAGNACAGTTGGAACAGAGGAAT-ATGGGCCAAACAGGATATCTGTGGTAAGCAGATTCTCT 785
Db 318 AAGAACAGATGAGACAGCTAGTGTATGGGGCCAAACAGGATATCTGTGGTAAGCAGATTCTCT 377

QY 786 GCCCGGCTCAGGGCCAAAGAACAGATGGTCCCGCAGATCGGTCGCCGCTCAGCAGATTTC 845
Db 378 GCCCGGCTCGGGCCAAAGAACAGATGGTCCCGCAGATCGGTCGCCGCTCAGCAGATTTC 437

QY 846 TAGAGAAACCATCAGATGTTTCCAGGGTGCCCCAAAGGACCTG-AAATGACCCCTGTGCCTTA 904
Db 438 TAGTGAATCATCAGATGTTTCCAGGGTGCCCCAAAGGACCTGAAATGACCCCTGTACCTTA 497

QY 905 TTTGAACATAACCAATCATGTTGCTTCTGCTCTTCTGCTGCGCGCTTCTGCTCCCCGAGCT 964
Db 498 TTTGAACATAACCAATCAGTTGCGTTCCTGCTTCTGCTGCGCGCTTCTGCTCCCCGAGCT 557

QY 965 CAATAAAGAGCCCCACACACCCCTCACTCGGGCGGCCAGTCTCCCGATAGACTGGTGCCTGCC 1024
Db 558 CAATAAAGAGCCCCACACACCCCTCACTCGGGCGGCCAGTCTCCCGATAGACTGGTGCCTGCC 617

QY 1025 CGGGTACCCGTTATTTCCCAATAAAGCCCTCTTGCTCTGTTTGCATCCGAAATCGTGGACTCGCTG 1084
Db 618 CGGGTACCCGTTATTTCCCAATAAAGCCCTCTTGCTCTGTTTGCATCCGAAATCGTGGACTCGCTG 677

QY 1085 ATCTTCGGAGGGGTCTCTCTCAGATTGATGATCTGCCCACTCGGGGGTCTTTTCATTTGGA 1144
Db 678 TTCCTTCGGAGGGGTCTCTCTCAGATTGATGATCTGCCCACTCGGGGGTCTTTTCATTTGGA 737

```

Qy	2222	CCGGTGGTACCTCACCCCTTHACAGTTCGGCGACACAGTGTGGGTCCGCGACACACAGACT	2288
Db	1734	CAGACAATCGGCTGTCTGTATGCCGCGTGTTCGGCTGTTCAGCGAGGGCGCCCGGTT	1793
Qy	2282	AAGAACTAGAACTCGCTGCGTGAAGGACCTTACACAGTCTCTGTGACCAACCCCAACCGCC	2341
Db	1794	CTTTTGTCAAGACCGACCTGTCCGGTGCCTGAAAGAACTGC-----	1836
Qy	2342	CTCAAAAGTAGACGGCATTCGAGCTTGGATACACGCCGCCCAACGTGAAGGCTGCCGACCCC	2401
Db	1837	-----AGGACGAGCGCGGGCTATCGTGGCTGGCCACGACGGCGTTCCTTTGCGCAGC	1891
Qy	2402	GGGGTGGACCATCTCTTAGACTGCCANTGGATGGAGCTGTATCATCTCTTCTTGGTAG	2461
Db	1892	TGTGCTCGACGTTGTCACTGAACGGGGAAGGAGCTGTGCTGTATTGGGGGAAGTGGCGGG	1951
Qy	2462	CAACAGCTACAGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTCAGCG	2521
Db	1952	GCAGGATCTCTGTCACTCACCTTGTCTCTCGCGAAGAAAGTATCCATCATNGTCTGATGC	2011
Qy	2522	CCAGCTGGGTGACAGAGTACCATCACCTGTAAAGCCAGTCAAGATGTGGGTACTTCTG	2581
Db	2012	AATCGCGGCGTGCATACGCTTGTATCGGCTACTTGCCTCATTCGACCAACGGAACA	2071
Qy	2582	TAGCTTGGTACACGAGAACCGAGTAAGGCTCAAAGCTGTGATCTACTTGACATPCCA	2641
Db	2072	TCGCATCGACGAGCAGCTACTCGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGA	2131
Qy	2642	CCGGCACACTGGTGTGTCGAAGCAGATTACGCGGTAGCGGTAGGGGTACCGACTTCACCT	2701
Db	2132	CGAAGAGCATCAGGGGCTCGCGCCAGCCGAAGTGT-----TCGCAGGC	2175
Qy	2702	TCACCATCAGCAGCCTCCAGCAGAGAGCATCGCCACTACTACTGTCGACCAATATAGCC	2761
Db	2176	TCAAGCGCGCATGCCCGACGCGAGGATCTCTGTCGTGACCAATGGCGATGCTCTTGC	2235
Qy	2762	TCTATCGGTCGTTGGCCAAAGGACCAAGGTGGAAATCAAACGAGGTGGCTCAGGATCGG	2821
Db	2236	CGAATATCATGTGTGAAATATGGCGCTTTTCTTGATTCATCGACTGTGGCGCGTGGGTG	2295
Qy	2822	GTGATCCGGCTCTGGTGGCTCAGGATCGAGGTTCCAACTGGTGGAGAGCGGTGGAGGTG	2881
Db	2296	TGGCGGACCGCTATCAGGACATACGCTTGGCTACCCGTGATATGTGTGAAGAGCTTGGCG	2355
Qy	2882	TTGTGCAACTGGCCGCTCCCTGGCGCTGCTCTGCCATCTGGCTTCGATTTCCACCA	2941
Db	2356	CGGAATGGGCTGACCGCTTCTCTGTGCTTTACGGTATCGCGCTCCC-----	2402
Qy	2942	CATATTGGATGAGTGGGTGAGACAGCACCTCGGAAAGGTCTTGAAGTGGATGGAGAAA	3001
Db	2403	-----GATTGCGAGCGCATCGCCTTCTATCGCTTCTTGACGAG	2441
Qy	3002	TTCTATCAGATAGCAGTACGATTAACTATGGCGCTCTTAAGGATAGATTTTACAATAT	3061
Db	2442	TTCTTCTGAGCGGACTCTGGGGTTGAAATGACCCGACGAGCGCAACCTGTCAT	2501
Qy	3062	CGCGAGACAGCCCAAGAACACATTGTTCTGTCAAAATGGACAGCCTCAGACCCGGAAGACA	3121
Db	2502	CACAGATTTTCGATTCACCGCCGCTTCTATGAAGGTTGGGCTTCGGAATCGTTTTCC	2561
Qy	3122	CCGGGTCTATTTTTGTGCAAGCCTTAACTTCGGCTTCCCTGGTGTGCTTATTTGGGGCC	3181
Db	2562	GGGACGC-----GGCTGGATGATCCT	2583
Qy	3182	AAGGGAACCCGGTCAACGCTCTCAGTGTCTAAGCCCAACGACGCCAGCGCGGACAC	3241
Db	2584	CCAGCGGGGATCTCATGTGTGAGTCTTTCGGCCACCCCGGGTTCGATCCCTCTCGCGAG	2643
Qy	3242	CAACACCGGCGCCACCATCGCGTCGAGCCCTGTCCCTGCGCCAGAGCGGCTCGGC	3301
Db	2644	TTGTTTCACTGTGCTTGAGGCTGAGACACTCG-----	2678
Qy	3302	CAGCGGGGGGGCGCAGTGACACGAGGGGGCTGGACTTTCGCCCTGGATCCCAAACTCT	3361

Db	2679	-----CGAGTTCTACCGCAGTGAATCCG	2705
Qy	3362	GCTACCTGCTGGATGGAATCCTTTCATCTATGGTGTCA	3421
Db	2706	TGGCATCCAGGA---AACCAGCAGCGGCTATCGGCGCATCCATGCCCCGAACTGCAGG	2762
Qy	3422	GAGTGAAGTTTCAGCAGAGGCGAGAGCCCCCGGTACACGACGAGGCGCAGAACCACTCT	3481
Db	2763	AGTGGGAGGACGATGGCCGCTTGGTGCACCCGACGGGACGCTCTCGCGCTGATAC	2822
Qy	3482	ATAACGAGCTCAATCTAGGACGAAGAGGAGGTACGATGTTTGGACAAGAGACGTGGCC	3541
Db	2823	AGAACGAATTCCTTTCAGGCGATCTCATGATGTGTCTT-----CCGTTTTTC	2869
Qy	3542	GGGACCTCGAGATGGGGGGAAGCCGAGAAAGGAAGAACCTCTCAGGAAGGCTCTCAATG	3601
Db	2870	CGCTGAGGTCACTGCGTGGATGGAGCGCTGGCGCTGTCTGCGCAGCGGCGAG---CTG	2925
Qy	3602	AATGCAAGAAGATAAGATGGGAGGCTTACAGTGAAGATGGGATGAAGAGGAGGCGCC	3661
Db	2926	CTCACCAACCACCTCGAGGGGTGTCAGCGCTGTCAGAGCGCGAGTGCAGAACTGC-----	2978
Qy	3662	GGAGGCGAAGGGCACGATGGCTTTTACAGAGGTCTCAGTACAGCCACCAGACACCT	3721
Db	2979	-----TCCAAAGGGACCTCAAGGCTTTCCAGGGGACACTAGGTGATCTCCATCGAGCCAG	3033
Qy	3722	ACGACGCGCTTCACATCAGCGCCCTGCGCCCTCGCTAACTCGACGGCGCGCGGATCCGG	3781
Db	3034	TGTA-----GAGATAAGCTTATCG	3052
Qy	3782	ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACA	3841
Db	3053	ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACA	3112
Qy	3842	ATATCACAGCTGAAGCCTATAGNATACGAGCCATAGATAAAATAAAGATTTTATTAG	3901
Db	3113	ATATCACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTAG	3172
Qy	3902	TCTCCAGAAAAAGGGGGAATGAAGACCCCACTCTGTAGTGTGGCAAGCTAGCTTAAGT	3961
Db	3173	TCTCCAGAAAAAGGGGGAATGAAGACCCCACTCTGTAGTGTGGCAAGCTAGCTTAAGT	3232
Qy	3962	AACGCCATTTTTCGAAGGCATGG-AAAATAATAACTGAGAAATGAGAAAGTTTCAGATCAAG	4020
Db	3233	AACGCCATTTTTCGAAGGCATGGAAAAATACATAACTGAGAAATGAGAAAGTTTCAGATCAAG	3292
Qy	4021	GTTAGGAACAGA-CAGACAGAGAAATATGGCCAAACAGGATATCTGTGTATAGCAGTTC	4079
Db	3293	GTCAGGAACAGATGGAAACAGCTGAATATGGGCCAAACAGAGATATCTGTGTATAGCAGTTC	3352
Qy	4080	CTGCCCC-GCTCAGGGCCAGAACAGTTGGAAACAGGAGAAATATGGGCCAAACAGGATATC	4138
Db	3353	CTGCCCCGCTCAGGGCCAGAACAGATGAAACAGCTGAATATGGGCCAAACAGGATATC	3412
Qy	4139	TGTGTTAAGCAGTTTCTGCCCCGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCGGTC	4198
Db	3413	TGTGTTAAGCAGTTTCTGCCCCGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCGGTC	3472
Qy	4199	CGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAA	4258
Db	3473	CAGGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAGGACCTGAA	3532
Qy	4259	ATGACCTGTGCTTATTTGAACATAACATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGC	4318
Db	3533	ATGACCTGTGCTTATTTGAACATAACATCAGTTCGCTTCTCGCTTCTGTTTCGCGCGC	3592
Qy	4319	TTCTGCTCCCCGAGCTCAATAAAGAGCCCAACCCCTCCTCGGCGCGCCAGTCTCTCC	4378
Db	3593	TTCTGCTCCCCGAGCTCAATAAAGAGCCCAACCCCTCCTCGGCGCGCCAGTCTCTCC	3652
Qy	4379	GATAGATGCGTTCGCGCGGTACCGGTGTTCTCAATAAAACCTCTTTCGAGTTGATCCGA	4438

6597 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTGTGCAAGTAAGT 6656
 5481 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTGTGCAAGTAAGT 5540
 6657 TGGCCGCGAGTGTATCACTCATGTTATGCGAGCAGCTGATATTTCTTACTGTCTATGC 6716
 5541 TGGCCGCGAGTGTATCACTCATGTTATGCGAGCAGCTGATATTTCTTACTGTCTATGC 5600
 6717 CATCGTAAAGTCTTTCTGACGTGGTGTGAGTACTCAACCAAGTCAATTCGAGAATAGT 6776
 5601 CATCGTAAAGTCTTTCTGACGTGGTGTGAGTACTCAACCAAGTCAATTCGAGAATAGT 5660
 6777 GTATCGCGACCGAGTGTCTCTTTCGGCGCGCTCAATACGGGATAATACCGGCCACATA 6836
 5661 GTATCGCGACCGAGTGTCTCTTTCGGCGCGCTCAACCGGATAATACCGGCCACATA 5720
 6837 GCAGAACTTTAAAGTGTCTCATCTTGGGAAAGCTTTTCGGGGGGAAGAACTCTCAAGGA 6896
 5721 GCAGAACTTTAAAGTGTCTCATCTTGGGAAAGCTTTTCGGGGGGAAGAACTCTCAAGGA 5780
 6897 TCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTGATCTTCAG 6956
 5781 TCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCCACTGATCTTCAG 5840
 6957 CATCTTTTACTTTTACACGCGTTCCTGCGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 7016
 5841 CATCTTTTACTTTTACACGCGTTCCTGCGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 5900
 7017 AAAAGGAATAAGGCGGACACGGAATGTTGAATCTATCTCTTCTTTTCAATATT 7076
 5901 AAAAGGAATAAGGCGGACACGGAATGTTGAATCTATCTCTTCTTTTCAATATT 5960
 7077 ATTGAAGCATTTATCAGGTTATTGCTCATGCGGATACATATTTGAATCTATTAGA 7136
 5961 ATTGAAGCATTTATCAGGTTATTGCTCATGCGGATACATATTTGAATCTATTAGA 6020
 7137 AAAATAAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG 7196
 6021 AAAATAAAACAATAGGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGCTTAAG 6080
 7197 AAACCATTAATATGACATTAACCTATAAAAATAGGCGTATACGAGGCCCTTTTCGTC 7256
 6081 AAACCATTAATATGACATTAACCTATAAAAATAGGCGTATACGAGGCCCTTTTCGTC 6140
 7257 T 7257
 6141 T 6141

RESULT 8
 US-08-848-760B-1
 ; Sequence 1, Application US/08848760B
 ; Patent No. 6248721
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Lung-Ji
 ; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: Florida
 ; COUNTRY: United States of America
 ; ZIP: 32606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/848,760B
 ; FILING DATE: 25-Jan-2001
 ; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/838,702
 FILING DATE: 09-APR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: PACE, DORAN R.
 REGISTRATION NUMBER: 38,261
 REFERENCE/DOCKET NUMBER: CNG-100C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (352) 375-8100
 TELEFAX: (352) 372-5800
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6145 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-08-848-760B-1
 Query Match 40.0%; Score 3062.2; DB 3; Length 6145;
 Best Local Similarity 71.7%; Pred. No. 0;
 Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;
 QY 549 TTTTGAAGAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAAGTAACGCCATTTTCAAGG 608
 DB 144 TTTTGAAGAGACCCACCCCGTAGG--TGGCAAGCTAGCTTAAAGTAACGCCACCTTTGCAAGG 201
 QY 609 CATGG-AAAATACATACTGAGATAGAGAGATTGAGATCAAGTTAGGAAACAGAGAGAC 667
 DB 202 CATGAAAAATACATACTGAGATAGAAAAAGTTGAGATCAAGGTCAGGAAACAAAGAAAC 261
 QY 668 AGCAGAAATATGGCCCAACACAGATATCTGTGTAGACAGTTCCTGCCCC-GCTCAGGGCC 726
 DB 262 AGCTGAATA-CCAAACAGAGATATCTGTGTAGACAGTTCCTGCCCCGGCTCAGGGCC 317
 QY 727 AAGAACAGTTGGAAACAGAGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAGACAGTTCCT 785
 DB 318 AAGAACAGATGAGACAGCTGAGTGATGGCCCAACAGGATATCTGTGTAGACAGTTCCT 377
 QY 786 GCCCGGCTCAGGGCCAAAGAACAGATGTCGCCAGATGGGTCCCGCTCCAGAGATTTC 845
 DB 378 GCCCGGCTCAGGGCCAAAGAACAGATGTCGCCAGATGGGTCCCGCTCCAGAGATTTC 437
 QY 846 TAGAACAATCATCAGATGTTTCCAGGGTCCGCCAAGGACCTG-AAATGACCTGTGCCCTTA 904
 DB 438 TAGTGAATCATCAGATGTTTCCAGGGTCCGCCAAGGACCTGAAAATGACCTGTGCCCTTA 497
 QY 905 TTTGAATCAACCAATCAGTTCGCTTCTGCTTCTGCTGCGCGCTTCTGTCTCCCGAGCT 964
 DB 498 TTTGAATCAACCAATCAGTTCGCTTCTGCTTCTGCTTCTGCTGCGCGCTTCTGTCTCCGAGCT 557
 QY 965 CAATAAAGAGAGCCCAACCCCTCACTCGGCGCGCAGTCTCCGATAGACTGCTCGCC 1024
 DB 558 CAATAAAGAGAGCCCAACCCCTCACTCGGCGCGCAGTCTTCCGATAGACTGCTCGCC 617
 QY 1025 CGGTACCCGTATTCCTCAATAAAGCTCTTGTGTGTTGATCCGATCCGATCGTGCAGTCTG 1084
 DB 618 CGGTACCCGTATTCCTCAATAAAGCTCTTGTGTGTTGATCCGATCCGATCGTGCAGTCTG 677
 QY 1085 ATCTTTGGGAGGCTCTCTCAGATGATGATGCTGCTGCGGACCTCGGGGGTCTTTTATTGGA 1144
 DB 678 TTCTTTGGGAGGCTCTCTCAGATGATGATGCTGCTGCGGAGGCTTTTATTGGA 737
 QY 1145 GGTTCACCGAGATTTGAGACCCCTGCGCCAGGAGCAACCGACCCCGCGCGGAGTA 1204
 DB 738 GGTTCGTCGGGATTTGAGACCCCTGCGCCAGGAGCAACCGA-CCACCAACCGCGGAGTA 796
 QY 1205 AGTGGCAGCAACTTATCTGTGTCTGTCGATGTTGTCTAGTGTCTTATGATGATTTATG 1264
 DB 797 AGTGGCAGCAACTTATCTGTGTCTGTCGATGTTGTCTAGTGTCTTATGATGATTTATG 856
 QY 1265 GCCTGTGCGTACTAGTAACTAGCTCTGATCTGTATCTGCGGAGCCCGTGGTGAAC 1324

[illegible]

1837	Db	-----AGGA	CGAGCGACGGCGGCTATCGTGGCTGGGCCACGACGAGCGCGCTTCTCTTGGCGAGC	1891
2402	Qy	GGGGGTGGACCATCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTTCTTGGTAG	2461	
1892	Db	TGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGAAGTGCCGGG	1951	
2462	Qy	CAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCGCTGAGCG	2521	
1952	Db	GCAGGATCTCTGTCTATCTCACTTGTCTCTCGCGAAGATTCATCATGCTGCTATGC	2011	
2522	Qy	CCAGCTGGGTGACAGAGTGACCATCACCTGTGAAGCCAGTCAAGTATGGGTACTTCTG	2581	
2012	Db	ATGCGCGCGCTGCATACGCTTGATCCGGCTACCTGCCAATTCGACCAACGCGAAACA	2071	
2582	Qy	TAGCTTGGTACCAGCAGAAAGCCAGTAAGGCTCCAAAGCTGCTGATCTACTGGACATCCA	2641	
2072	Db	TCGCATCGAGGAGCAGTACTCGGATGGGAAGCGGCTTGTGATCAGAGTATCTGGA	2131	
2642	Qy	CCCGGCACACTGGTGTGCAAGCAGATTACGCGGTAGCGGTACGGGTACCGATTTCACCT	2701	
2132	Db	CGAAGAGATCAGGGGGTCGCGCCAGCCGAATGTT-----TCGCCAGGC	2175	
2702	Qy	TCACCATCAGCAGCTCCAGCCAGAGGACATCCGCACTACTACTGCCAGCATATAGCC	2761	
2176	Db	TCAAGCGCGCATGCCGACGCGGAGGATCTCGTCGTGACCCATGGCGATGCTGCTGTGC	2235	
2762	Qy	TCATCGGTGTTTCGGCCAAAGGACCAAGGTGAAATCAAAACGAGGTGGCTCAGGATCGG	2821	
2236	Db	CGAATATCATGTTGGAAATGGCGCGCTTTTCTGGAATCATCGACTGTGGCCGGCTGGGTG	2295	
2822	Qy	GTGGATCCGGCTCTGGTGGCTCAGGATCGAGGTTCCAACTTGGTGGAGAGCGGTGGAGGTG	2881	
2296	Db	TGGCGGACCCGCTATCAGGACATAGCTGTGGCTACCCGCTGATATTGCTGAAGAGCTTGGCG	2355	
2882	Qy	TTGTGCAACTGCGCGGTCCTCGGCTGTCTGCTCCGCATCTGGCTTCGATTTTACCA	2941	
2356	Db	GCGAATGGGCTGACCGGCTTCTCGTGTCTTACGGGTATCGCGCTGCC-----	2402	
2942	Qy	CATATTGGATGAGTTGGGTGAGACAGGCACCTCGGAAAGGTCTTTGAGTGGATTGGAGAAA	3001	
2403	Db	-----GATT	CGACGCATCGCCTTCTATCGCCTTCTTGACGAG	2441
3002	Qy	TTCATCCAGATAGCAGTACGATTAACTATGCGCCGCTCTTAAAGGATAGATTTACAATAT	3061	
2442	Db	TTCTTCTGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCAT	2501	
3062	Qy	CGCGAGAACACCGCAAGAACATTTGTTCTGCAAAATGACAGCCTGAGACCCGAAAGACA	3121	
2502	Db	CAGGAGATTTTCGATTCCACCGCGCGCTTCTATGAAAGTTTGGGCTTCGGAATCGTTTTC	2561	
3122	Qy	CCGGGTGCTATTTTGTGCAAGCCTTTACTTTCGGCTTCCCTGGTTTGTCTATTGGGGCC	3181	
2562	Db	GGGACGCC-----GGCT	TGGATGATCTCT	2583
3182	Qy	AAGGGACCCCGGTACCGCTTCCAGTGTCAAGCCCAACGACGCGCAGGCGCCCGCAC	3241	
2584	Db	CCAGCGGGGATCTCATGCTGGAGTTCTTCGCCCAACCCGGGCTCGATCCCTCGGAG	2643	
3242	Qy	CACACCGGCGCCACCATGCGGTGCGAGCCCTGTCTTCCTGCGGCCAGAGCGGCTCGGC	3301	
2644	Db	TTGTTTCAGCTGTGCTGAGGCTGACGACCTCG-----	2678	
3302	Qy	CAGCGGGGGGCGAGTGCCACACGAGGGGGTGGACTTCGCCCTGGATCCCAACTCT	3361	
2679	Db	-----CGGAGTTCTACCGGCAGTGCAAATCCG	2705	
3362	Qy	GCTACCTGTGTGATGGAATCCTTTCATCTATGGTGTCAATTTCTCATGTGCTTGTCTCTGA	3421	
2706	Db	TCGGCATCCAGGA--AAC	CAGCAGCGGCTATCCGGCATCCATGCCCCCGAACTCGAGG	2762
3422	Qy	GAGTGAAGTTTCAGCAGGACGCGACAGACCCCGCTGTCAGCAGAGGCGCAAGACCGCTCT	3481	
2763	Db	AGTGGGAGAGCAGATGGCGCTTTGGTTCGACCCGAGCGGAGCGCTCTTGGCGCTGATAC	2822	

QY	3482	ATAACGAGCTCAATCTAGACGAAGAGAGAGAGTA	CGATGTTTGTGACAAAGAGACGTTGGCC	3541
DB	2823	AGAACGAATGCTTGACGGCATCTCATGAGTGTCTT	-----CCGCTTTTC	2869
QY	3542	GGGACCTGAGATGGGGGGAAGCCGAGAGGAAGAAC	CCCTCAGGAAGGCCTGTACATATG	3601
DB	2870	CGCCTGAGGTCACTGCGTGGATGGAGCGCTGGCGCT	GTGCGCAGCGCGAG-----CTG	2925
QY	3602	AACTGCAGAAAGATAAGATGGCGGAGCCCTACAGT	GAGATTGGGATGAAGCGCAGCGCC	3661
DB	2926	CTCACCACCACTCGAGGGCGTGCAGCGCTGCAGAG	CGCGAGTGCAGAACTGTC-----	2978
QY	3662	GGAGGGCAAGGGCAAGATGGCTTTTACAGGGTCT	CAGTACAGCCACCAAGAGCACCT	3721
DB	2979	-----TCCAAAGGGAACCTCAAGGCTTTCCGAGG	GACACTAGGCTGACTCCATCAGGCCAG	3033
QY	3722	ACGAGCCCTTCACATGCAGGGCCCTGCCCTTCG	CTTAACCTCGACGCGCCGCGATCCGG	3781
DB	3034	TGTA-----	-----GAGATAAGCTTATCG	3052
QY	3782	ATTAGTCCAATTTGTTAAAGACAGGATATCAGT	TGTCAGGCTCTAGTTTTGACTCAACA	3841
DB	3053	ATTAGTCCAATTTGTTAAAGACAGGATATCAGT	TGTCAGGCTCTAGTTTTGACTCAACA	3112
QY	3842	ATATCACCAGCTGAAGCCCTATAGAGTACGAGC	CATAGATAAAATAAAGATTTTATTAG	3901
DB	3113	ATATCACCAGCTGAAGCCCTATAGAGTACGAGC	CATAGATAAAATAAAGATTTTATTAG	3172
QY	3902	TTCTCCAGAAAAGGGGGGAATGAAGACCCCACT	GTAGTTTGGCAAGCTAGCTTAAGT	3961
DB	3173	TTCTCCAGAAAAGGGGGGAATGAAGACCCCACT	GTAGTTTGGCAAGCTAGCTTAAGT	3232
QY	3962	AAGCCATTTTGCAGGCGATGG-AAAATACATA	CTGAGATAGAGAGTTTCAGATCAAG	4020
DB	3233	AAGCCATTTTGCAGGCGATGGAAATAACATACT	GAATAAGAGAGTTTCAGATCAAG	3292
QY	4021	GTTAGGAACAGA-GAGACAGCAGAATATGGG	CCAAACAGGATATCTGTGTAAACAGTTC	4079
DB	3293	GTCAAGNACAGATGGAAACAGCTGATATGGG	CCAAACAGGATATCTGTGTAAACAGTTC	3352
QY	4080	CTGCCCC-GCTCAGGCGCCAAAGACAGTTGG	AAACAGAGAAATATGGGCCAAACAGGATATC	4138
DB	3353	CTGCCCCGCTCAGGCGCCAAAGACAGATGGA	ACAGCTGAATATGGGCCAAACAGGATATC	3412
QY	4139	TGTGTTAAGCAGTTCTGCCCCGGCTCAGGG	CCAAAGACAGATGCTCCCCCAGATGCGGTC	4198
DB	3413	TGTGTTAAGCAGTTCTGCCCCGGCTCAGGG	CCAAAGACAGATGCTCCCCCAGATGCGGTC	3472
QY	4199	CCGCCCTCAGCAGTTCTTAGAGAACCATCAG	ATGTTTCCAGGGTGCCCCAAAGCACTTGAA	4258
DB	3473	CAGCCCTCAGCAGTTCTTAGAGAACCATCAG	ATGTTTCCAGGGTGCCCCAAAGCACTTGAA	3532
QY	4259	ATGACCTGTGCTTATTTGAATTAACCAATC	AGTTCTGCTTCTGTTCTGCGCGC	4318
DB	3533	ATGACCTGTGCTTATTTGAATTAACCAATC	AGTTCTGCTTCTGTTCTGCGCGC	3592
QY	4319	TTCTGCTCCCGAGCTCAATAAAGAGCCCA	CAACCCCTCAGTGGGCGCGCAGTCTCC	4378
DB	3593	TTCTGCTCCCGAGCTCAATAAAGAGCCCA	CAACCCCTCAGTGGGCGCGCAGTCTCC	3652
QY	4379	GATAGACTGCTGCGCGCGGTACCGGTGTTCT	CAATAAACCCTCTGCAGTTGATCCGA	4438
DB	3653	GATAGACTGCTGCGCGCGGTACCGGTGTTCT	CAATAAACCCTCTGCAGTTGATCCGA	3712
QY	4439	CTCGTGGTCTCGCTGTTCTTGGAGGGTCT-	CTCTGAGTGATTGACTTACCCGTCAGCGG	4497
DB	3713	CTTGTGCTCTCGCTGTTCTTGGAGGGTCTCC	TCTGAGTGATTGACTTACCCGTCAGCGG	3772
QY	4498	GGTCTTTTCAAGTTTCTCCCAACCTTACAG	GTCTCAATTTCTGTGATGTGCGCGAGGGA	4557
DB	3773	GGGTCTTTTCAATTTTGGGGGCTCTGTCGGG	ATCGGGAGACCCCTTGCCCAAGGACCCAGC	3832
QY	4558	CTCCGTCAGCCGGTTTTTGTATTATAAATA	TCCAAGACAGTGTTCCTTCAAGCCAG	4611
DB	3833	CACCAACCGGAGGTAAAGCTG-	-----	3851
QY	4618	ACTACATCTGACTCTCGGCTTTATAAAGA	ATCTGTAAGGGCTCTGTGACTATCTGCC	4677
DB	3853	-----GCTGCTCGCCGTTTTCGGTGATG	ACGGTGAACCTCTGAC-ACATGACGT	3904
QY	4678	ACACGACTTTTTAAGATTTTTATGCCTCT	CGATGAGGGATTTAGTCAATCTATCTCTGT	4737
DB	3905	CCCGAGACGGTTCACAGCTTCTGCTAAG	CGGATGCCGGAGCAGACAAGCCGTCAGGG	3964
QY	4738	CTATTTTGTGTGCTCTCTCCGATTTTAAAT	TTTCTAGTTTGACATCTCCCTCTGAGAGCAC	4797
DB	3965	CGCGTCAGCGGGTGTGGCG-	-----	3984
QY	4798	GCGGATTGCAGAGTAGTTAATACTCTGAG	GGCAGGCTTCTGTGAAAAGTTGCTGCGCT	4857
DB	3985	-----	-----	3984
QY	4858	CAGTGTGAGATTTTGGCCATAAAAGGGGT	CTCTGCCCTGTGTATACAGACAGATCGGAATCT	4917
DB	3985	-----	-----	3984
QY	4918	AGAGTGCAATCTCAGAGTCCCGCGGTTCC	GGGGCTCTGATCTCAGGGCAATCTTTGCTTA	4977
DB	3985	-----GGTGTGGGGCGCAGCATGACCCAGTC-	-----	4012
QY	4978	GAGATCCTCTACGCCGGAGCGCATCGTGG	CCGGGTACCGAGCTCGAATTCGTAATCATGGT	5037
DB	4013	-----ACGTAGCGATAGCGGAGTGTAT	CTGCTGCTTAATTCGCGCAT	4055
QY	5038	CATAGCTGTTTCTGTGTGAAAATTTGTTAT	TCGCTCACAATTTCCACAACAATACGAGCCG	5097
DB	4056	CAGACAGATTTGACTGAGAG-	-----	4076
QY	5098	GAAGCATAAAGTGTAAAGCCCTGGGGT	GCCTAAATGAGTGAGCTAACTCACATTAATTCGCT	5157
DB	4077	-----	-----	4076
QY	5158	TGCGTCTACTGCCCGCTTTCCAGTCTGG	AAACCTGTCTGCCAGCTGCATTAATGAATCG	5217
DB	4077	-----TGCAACCATATGCGGTGTGAATATC	-----	4100
QY	5218	GCCAAACGCGGGGAGAGCGGTTTTGCGT	ATTGGCGCTCTTCGCTTCTCGCTCACTG	5277
DB	4101	CGCACAGATGCTTAAGGAAAAATACCG	CATCAGCGCTCTTCGCTTCTCGCTCACTG	4160
QY	5278	ACTCTGCTCGCTCGGCTCGGCTCGGG	CGAGCGGTATCAGCTCACTCAAAAGCGGTAA	5337
DB	4161	ACTCTGCTCGCTCGGCTCGGCTCGGG	CGAGCGGTATCAGCTCACTCAAAAGCGGTAA	4220
QY	533			

Db 4521 CACGCTGAGGTATCTCAGTTTCGGTGTAGGTCGTTTCGCTCAAGCTGGGCTGTGTGCACG 4580
Qy 5698 AACCCCGGTTTACGCGGACCGCTCGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACC 5757
Db 4581 AACCCCGGTTTACGCGGACCGCTCGCGCTTATCCGGTAACTATCGTCTTGAGTCCAAACC 4640
Qy 5758 CGGTAAAGACAGCACTTATCGGCACTGGGACAGGACCACTGGTAAACAGGATTAAGCAGAGCA 5817
Db 4641 CGGTAAAGACAGCACTTATCGGCACTGGGACAGGACCACTGGTAAACAGGATTAAGCAGAGCA 4700
Qy 5818 GGTATGAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACCGCTACACTAGAA 5877
Db 4701 GGTATGAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACCGCTACACTAGAA 4760
Qy 5878 GCACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTTACCTTCGGAAAAAGAGTTGGTA 5937
Db 4761 GCACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTTACCTTCGGAAAAAGAGTTGGTA 4820
Qy 5938 GCTCTTGATCGGCAAAACAAACCCAGCTGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGC 5997
Db 4821 GCTCTTGATCGGCAAAACAAACCCAGCTGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGC 4880
Qy 5998 AGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGCTCG 6057
Db 4881 AGATTACGCGCAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGCTCG 4940
Qy 6058 ACAGTACAGTGAACGAAACCTCACGTTAAAGGATTTGGTCAATGAGATTATCAAAAGGA 6117
Db 4941 ACAGTACAGTGAACGAAACCTCACGTTAAAGGATTTGGTCAATGAGATTATCAAAAGGA 5000
Qy 6118 TCTTCACTAGATCCTTTTAAATTAATAAGTATTAATCAATCTAAAGTATATATG 6177
Db 5001 TCTTCACTAGATCCTTTTAAATTAATAAGTATTAATCAATCTAAAGTATATATG 5060
Qy 6178 AGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGGATCT 6237
Db 5061 AGTAAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGACCTATCTCAGGATCT 5120
Qy 6238 GTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGCTGATGATACTACGATACGGG 6297
Db 5121 GTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGTCTGCTGATGATACTACGATACGGG 5180
Qy 6298 AGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCAGCGGCTC 6357
Db 5181 AGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCAGCGGCTC 5240
Qy 6358 CAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGGAGGAGAGTGGTCTCTGCAA 6417
Db 5241 CAGATTTATCAGCAATAAACAGCAGCCGGAAGGCGGAGGAGAGTGGTCTCTGCAA 5300
Qy 6418 CTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGCTAGAGTAGTAGTTCGC 6477
Db 5301 CTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAGCTAGAGTAGTAGTTCGC 5360
Qy 6478 CAGTTAATAGTTTGGCAACGTTGTCATTTGCTACAGGC-TGGTGTGTACAGCTCGT 6536
Db 5361 CAGTTAATAGTTTGGCAACGTTGTCATTTGCTGAGGATCGTGTGTGTCAGCTCGT 5420
Qy 6537 CGTTTGGTATGGCTTCAATCAGCTCGGTTCCCAACGATCAAGCGAGTTTACATGATCCC 6596
Db 5421 CGTTTGGTATGGCTTCAATCAGCTCGGTTCCCAACGATCAAGCGAGTTTACATGATCCC 5480
Qy 6597 CCATGTTGTGCAAAAGCGGTAGCTCCTTCGGTCCCTCCATCGTTGTGAGAGTAGT 6656
Db 5481 CCATGTTGTGCAAAAGCGGTAGCTCCTTCGGTCCCTCCATCGTTGTGAGAGTAGT 5540
Qy 6657 TGGCGGAGTGTATCACTCAATGTTATGGCAGCACTGCAATAATCTCTTACTGTATGC 6716
Db 5541 TGGCGGAGTGTATCACTCAATGTTATGGCAGCACTGCAATAATCTCTTACTGTATGC 5600
Qy 6717 CATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGATAGT 6776

Db 5601 CATCCGTAAGATGCTTTTCTGTGACTGGTGTAGTACTCAACCAAGTCAATCTTGAGATAGT 5660
Qy 6777 GTATGCGGCGACCGAGTTGCTCTTTCGCCGGGTCAATACGGGATAATACCGCGCCACATA 6836
Db 5661 GTATGCGGCGACCGAGTTGCTCTTTCGCCGGGTCAACACGGGATAATACCGCGCCACATA 5720
Qy 6837 GCAGAACTTTAAAGTGTCTCATTTGGAAGAGTCTTTCGGGGCGAAAACTCTCAAGGA 6896
Db 5721 GCAGAACTTTAAAGTGTCTCATTTGGAAGAGTCTTTCGGGGCGAAAACTCTCAAGGA 5780
Qy 6897 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACCCAACTGATCTTCAG 6956
Db 5781 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCACCCAACTGATCTTCAG 5840
Qy 6957 CATCTTTTACTTTTACCAGCGCTTTCTGGGTGAGCAAAAAACAGGAGGCAAAATGCCGCAA 7016
Db 5841 CATCTTTTACTTTTACCAGCGCTTTCTGGGTGAGCAAAAAACAGGAGGCAAAATGCCGCAA 5900
Qy 7017 AAAAGGAGATAAGCGGCGACACGGAATTTGAATACTCATACTCTTCTCTTTTCAATATT 7076
Db 5901 AAAAGGAGATAAGCGGCGACACGGAATTTGAATACTCATACTCTTCTCTTTTCAATATT 5960
Qy 7077 ATTCAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTATTTAGA 7136
Db 5961 ATTGAAGCATTTATCAGGTTTATGCTCATGAGCGGATACATATTTGAATGTATTTAGA 6020
Qy 7137 AAAATAAACAAATAGGGGTTCCCGCGCACATTTCCCGAAAAAGTCCACTCACGCTCTAAG 7196
Db 6021 AAAATAAACAAATAGGGGTTCCCGCGCACATTTCCCGAAAAAGTCCACTCACGCTCTAAG 6080
Qy 7197 AAACCATTTATCATGACATTAACCTTAACTATAAAATAGCGGTATCACGAGGCCCTTTCTGC 7256
Db 6081 AAACCATTTATCATGACATTAACCTTAACTATAAAATAGCGGTATCACGAGGCCCTTTCTGC 6140
Qy 7257 T 7257
Db 6141 T 6141

RESULT 9
US-09-826-025-1
; Sequence 1, Application US/09826025
; Patent No. 6730512
; GENERAL INFORMATION:
; APPLICANT: Chang, Lung-Ji
; TITLE OF INVENTION: Combination Immunogene Therapy
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 04-Apr-2001
; APPLICATION NUMBER: US/09/826,025
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/838,702
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: CHANG-02687
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 1:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 6145 base pairs									
; TYPE: nucleic acid									
; STRANDEDNESS: double									
; TOPOLOGY: linear									
; MOLECULE TYPE: DNA (genomic)									
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:									
US-09-826-025-1									
Query Match 40.0%; Score 3062.2; DB 4; Length 6145;									
Best Local Similarity 71.7%; Pred. No. 0;									
Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;									
Qy	549	TTTTGAAAGACCCCACTGTAGGTTTCGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGG	608						
Db	144	TTTTGAAAGACCCCACTCGGTAGG--TGCGAAGCTAGCTTAAGTAACGCCACTTTCGAGG	201						
Qy	609	CATGG-AAAATACATAACTGAGATAGAGAAGTTTCAGATCAAGGTTTAGGAACAGAGAC	667						
Db	202	CATGGAAAAATACATACTGAGATAGAAAAGTTTCAGATCAAGGTCAGGAACAAGAAAC	261						
Qy	668	AGCAGATATGGGCCAAACAGGATATCTGTGTTAAGCAGTTCTCTGCCCC--GCTCAGGGCC	726						
Db	262	AGCTGGAAT-----CCAAACAGGATATCTGTGTTAAGCGGTTCTCTGCCCCCGGCTCAGGGCC	317						
Qy	727	AAGAACAGTTTGGAAACAGAGAAAT-ATGGGCCAAACAGGATATCTGTGTTAAGCAGATTCC	785						
Db	318	AAGAACAGATGAGACACTGATGATCGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC	377						
Qy	786	GCCCCGGCTCAGGGCCAAAGAACAGATGGTCCCCAGATGCGGTCCGCCCTCAGCAGTTTC	845						
Db	378	GCCCCGGCTCGGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCAGTTTC	437						
Qy	846	TAGAGAACCATCAGATGTTTCCAGGTTGCCCAAGGACCTTG-AAATGACCTGTGCCTTA	904						
Db	438	TAGTGAATCATCAGATGTTTCCAGGTTGCCCAAGGACCTGAAATGACCTGTGACCTTA	497						
Qy	905	TTTCAACTTAACCAATCAGTTTCGCTTTCGCTTCTGTTTCGGCGCTCTCTGCTCCCGAGCT	964						
Db	498	TTTGAACTTAACCAATCAGTTTCGCTTTCGCTTCTGTTTCGGCGCTTCCGCTCTCCGAGCT	557						
Qy	965	CAATAAAGAGCCCCAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGTCGCC	1024						
Db	558	CAATAAAGAGCCCCAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGCGTCGCC	617						
Qy	1025	CGGTAACCGTATTTCCAAATAAGACCTCTTGCTGTTTGCATCCGAATCGTGGACTGCGTG	1084						
Db	618	CGGTAACCGTATTTCCCAATAAAGCCCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTG	677						
Qy	1085	ATCCTTCGGAGGTCCTCTCAGATTGATTCAGCTGCCACCTCGGGGCTCTTCATTGGA	1144						
Db	678	TTCTCTGGAGGGTCTCTCTGAGTGAATGACTACCCACGACGGGGTCTTTCAATTGGG	737						
Qy	1145	GGTTTCCACCGAGATTGGAGACCCCTGCCCCAGGACCAACGACCCCCCGCCGGAGGTA	1204						
Db	738	GGCTCGTCGGGATTGGAGACCCCTGCCACGAGCAACCGA-CCACACACCGGAGGTA	796						
Qy	1205	AGCTGGCCAGCAACTTATCTGTGTCGTGTCGATTTGTTAGTGTCTATGACTGATTTTATG	1264						
Db	797	AGCTGGCCAGCAACTTATCTGTGTCGTGTCGATTTGTTAGTGTCTATGTTTGTGTTATG	856						
Qy	1265	CGCCTGCGTGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGTGGTGAAC	1324						
Db	857	CGCCTGCGTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCGTGGTGAAC	916						
Qy	1325	GACGAGTTTCGGAACAACCCGGCCGCAACCTCTGGAGAGCGTCCAGGAGACTTCGGGGCCGT	1384						
Db	917	GACGAGTTCTGAAACACCCGGCCGCAACCTCTGGAGAGCGTCCAGGAGACTTCGGGGCCGT	976						
Qy	1385	TTTTTGGCCCCGACTGAGTCCCTAAATCCGATTCGTTTATGACTCTTTTGTGTCACCCCC	1444						
Db	977	TTTTTGGCCCCGACTGAGGAGGGAGTTCATGTGGAAATCCGACCC-----	1022						

QY 2522 CCAGGTGGGTGACAGAGTGACCATCACTGTAAGGCCAGTCAGATGTGGTACTTCTG 2581
Db 2012 AATCGCGGGTGCATACGTTGATCCGGCTACCTGCCCATTCGACCAACGAAGCAACA 2071
QY 2582 TAGCTTGTGTACAGCAGAAAGCAGGTAAGGCTCCAAAGCTGCTGATCTACTGTGACATCCA 2641
Db 2072 TCGCATCGAGCGAGCAGTACTCGATCGAATGAAGCCGGTCTTGTGATCAGGATGATCTGGA 2131
QY 2642 CCGGCGACACTGTGTGTGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACCT 2701
Db 2132 CGAAGAGCATCAGGGGCTCGCGCCAGCCGAATGT-----TCGCCAGGC 2175
QY 2702 TCACCATCAGCAGCTCCAGCCAGAGGACATCGCCACCTACTACTGCCAGCAATATAGCC 2761
Db 2176 TCAAGGCGCATGCCCCAGCGGAGGATTCGTCGTGACCCATGGCGATGCCCTGTGC 2235
QY 2762 TCTATCGGTCTCGGCCAAGGGACCAAGGTGGAAATCAAAAGAGGTGGCTCAGGATCGG 2821
Db 2236 CGAATATCATGTTGGAAATGGCGCTTTCTGGATTCATCGACTGTGGCCGGCTGGGTG 2295
QY 2822 GTGATCCGGCTCTGTGGCTCAGGATCGGAGGTCCAACTGGTGAGAGCGGTGGAGTG 2881
Db 2296 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCCGTGATATTGCTGAAGAGCTTGGCG 2355
QY 2882 TTGTGCAACTGGCGGCTCCCTGCGCTGCTCTGCTCGCATCTGGCTTCGATTCACCA 2941
Db 2356 GCGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCC-----2402
QY 2942 CATATTGATGAGTTGGGTGAGACAGGACCTCGAAAGAGTCTTGAGTGGATTGGAGAA 3001
Db 2403 -----GATTCGAGCGCATCGCTTCTATCGCTTCTTGAACGAG 2441
QY 3002 TTCAATCCAGATAGCAGTACGATTAACATGATCGCGCTCTCTAAGGATAGATTTACAATAT 3061
Db 2442 TTCTTCTGAGCGGACTCTGGGGTTCGAAATGACCGACCAAGCAGCGCCCAACCTGCCAT 2501
QY 3062 CGCAGACAAACGCAAGNACATGTTCTCGCAATGGAAGCAGCTGAGACCCGAGACA 3121
Db 2502 CACGAGATTTTCGATTCACACCGCCGCTTCTATGAAGGTGGGGCTTCGGAATCGTTTTC 2561
QY 3122 CCGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGTTGCTTTATTTGGGCCC 3181
Db 2562 GGGAGGCC-----GGCTGGATGATCCT 2583
QY 3182 AAGGAGCCCCGGTCAACGTTCTCAGTGTCAAGCCACACGACGCGCAGCGCGGACCCAC 3241
Db 2584 CCAGCGGGGATCTCATGCTGGAGTCTTCGCCACCCCGGGCTCGATCCCTCGCGAG 2643
QY 3242 CAACACCGGCGCCACCATCGCGTGCAGCCCTGTCCTCGGCCAGAGCGGCTCGGC 3301
Db 2644 TTGGTTCAAGTGTGCTGCTGAGGCTGGACGACTCG-----2678
QY 3302 CAGCGGGGGGGCGCAGTGACACAGGGGGCTGAGCTCGCCCTCGATCCCAACTCT 3361
Db 2679 -----CGAGTTCTTACCGGAGTGCATAATCCG 2705
QY 3362 GCTACTGCTGGATGGAATCTCTTCAATATGTTGTCATCTCACTGCCCTTGTTCCTGA 3421
Db 2706 TCGGCATCCAGGA---AACCAGCAGCGCTATCCGCGATCCATGCCCGCCGAACTGCAGG 2762
QY 3422 GAGTGAAGTTGAGCAGAGCGGAGAGCCCCCGGGTACCGAGAGGCGCAGAACAGCTCT 3481
Db 2763 AGTGGGAGGACGATGCGCTTTGTGTGACCGCGAGCGGACGCTCTCGCGCTGTATAC 2822
QY 3482 ATACGAGCTCAATCTAGACGAGAGAGGAGTACGATGTTTGGACAAGAGAGTGGCC 3541
Db 2823 AGAAGCAATGCTGACGAGCATCTCATGAGTGTGTCTT-----CCCGTTTTTC 2869
QY 3542 GGGACCTTGAGATGGGGGAAAGCGAGAGGAAGAACCTTCAGGAAGGCGCTGTACAATG 3601
Db 2870 CGCCTGAGTCACTGCGTGGATGAGCGCTGCGGCTGCTGCGCGACGGCGAG---CTG 2925
QY 3602 AACTGCAGAAAGATAAGATGGCGAGGCGCTACAGTGAATTTGGGATGAAGGCGAGCGCC 3661

Db 2926 CTACACCCACCTCGAGGGCGTGCAGCGCTGCAGAGCCGAGTGCAGAACTGC-----2978
QY 3662 GGAGGGGCAAGGGGCAAGTGGCCCTTTTACAGGGTCTCAGTACAGCCACCAAGGACACT 3721
Db 2979 -----TCCAAAGGGAGCTCAAGGCTTTCCGAGGAGCACTAGGCTGACTCCATCGAGCCAG 3033
QY 3722 ACGAGCGCTTTCATCGAGCGCTCGCCCTCTCTAACTCGACGCGCGCGGATCCGG 3781
Db 3034 TGTA-----GAGATAAGCTTATCG 3052
QY 3782 ATTAGTCCAAATTTCTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTTGACTCAACA 3841
Db 3053 ATTAGTCCAAATTTCTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTTTGACTCAACA 3112
QY 3842 ATATCACCAGCTGAAGCTATAGAGTACGAGCCATAGATAAAATAAAAGATTTTATTTAG 3901
Db 3113 ATATCACCAGCTGAAGCTATAGAGTACGAGCCATAGATAAAATAAAAGATTTTATTTAG 3172
QY 3902 TCTCCAGAAAAGGGGGGAATGAAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAGT 3961
Db 3173 TCTCCAGAAAAGGGGGGAATGAAAGACCCACCTGTAGTGTGGCAAGCTAGCTTAAGT 3232
QY 3962 AACGCCATTTTGCAGGCATGG-AAAATACATACTCAGATAGAGAGTTTCAGATCAAG 4020
Db 3233 AACGCCATTTTGCAGGCATGGAAAATAACATACTCAGATAGAGAGTTTCAGATCAAG 3292
QY 4021 GTTAGGAACAGA-GAGACAGCAGATAATGGGCCAAACAGGATATCTGTGTAAAGCAGTTC 4079
Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAAAGCAGTTC 3352
QY 4080 CTGCCCC-GCTCAGGGCCAAAGAACAGTTGGAAACAGGAGATATGGGCCAAACAGGATATC 4138
Db 3353 CTGCCCCGGCTCAGGGCCAAAGAACAGATGGAACAGTGAATATGGGCCAAACAGGATATC 3412
QY 4139 TGTGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAACAGATGTTCCCAAGATGCGGTC 4198
Db 3413 TGTGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAACAGATGTTCCCAAGATGCGGTC 3472
QY 4199 CCGCCCTCAGCAGTTCCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAA 4258
Db 3473 CAGCCCTCAGCAGTTCCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCCAAGGACCTGAA 3532
QY 4259 ATGACCTGTGCTTATTTGAACTAACATCAGTTCGCTTCTCGTTCCTGTTCCGCGGC 4318
Db 3533 ATGACCTGTGCTTATTTGAACTAACATCAGTTCGCTTCTCGTTCCTGTTCCGCGGC 3592
QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCACACCCCTCAGTCCGCGCGCAGTCTCC 4378
Db 3593 TTCTGCTCCCGAGCTCAATAAAGAGCCACCAACCCCTCAGTCCGCGCGCAGTCTCC 3652
QY 4379 GATAGACTGCTGCTCCCGGTACCCGTGTTCTCAATAAACCCCTCTTGAGTTGCATCCGA 4438
Db 3653 GATTGACTGAGTGCCTCCGGGTACCCGTGTATCCAAATAAACCTCTTGAGTTGCATCCGA 3712
QY 4439 CTGTGTGCTCGCTGTTCTTGGAGGGTCT-CTCTGAGTGAATGACTACCCGTCAGCGG 4497
Db 3713 CTGTGTGCTCGCTGTTCTTGGAGGGTCTCTCTGAGTGAATGACTACCCGTCAGCGG 3772
QY 4498 GGTCTTTTCACTTCTCCCACTACACAGTCTCAATAACATCTCTGATGTGCGCAGGGA 4557
Db 3773 GGGTCTTTTCAATTTGGGGGCTCGTCCGGGATCGGAGACCCCTGCCAGGGACCCAGCC 3832
QY 4558 CTCCGTGAGCGCGGTTTTTTTATATAATAAATGAAGAACAGTGTTCCTTCAAGCCAG 4617
Db 3833 CACCACCGGAGGTAAAGCTG-----3852
QY 4618 ACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGGCTCTGTGAGTATCTGCC 4677
Db 3853 -----GCTGCTCGCGGCTTTTCGGTGTGACGGTGAAGAACTCTGAC-ACATGACGT 3904
QY 4678 ACAGACTTTTAAAGATTTTATGCCCTCTCGGTGAGGGATTTAGTCAATCTATCTCTGT 4737

3905 CCCGAGACGGTACACAGCTTGTCTGTAAAGCGGATCCGGGAGCAGACAAGCCCGTCAGGG 3964
4738 CTATTTGCTGCTCTCGTATTTAAATTTCTAGTTGCACTCCCTTCCTGAGAGCAC 4797
3965 CCGCTCAGCGGTGTGGCG----- 3984
4798 GCGGATGACAGAGTAGTTAATACTCTGAGGCGAGCGTCTGTGAAAAGGTGCTGGGCT 4857
3985 ----- 3984
4858 CAGGTGAGATTTTGCCATAAAAAAGGGTCTGTGCCCTCTGTACACAGACAGATCGGAATCT 4917
3985 ----- 3984
4918 AGAGTGCATACTCAGAGTCCCAGCGTTCGGGGCTCTGATCTCAGGGCATCTTTGCCTA 4977
3985 -----GGTGTGGGGCGCAGCCATGACCCAGTC----- 4012
4978 GAGATCCTCTAGCCGCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGGT 5037
4013 -----ACGTAGCAGTAGCGAGGTGATCTGGCTTAATATGGGCAT 4055
5038 CATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATCCACACACATACGAGCGG 5097
4056 CAGAGCAGATTGTACTGAGAG----- 4076
5098 GAAGCATAAAGTGTAAAGCTGGGTGCTTAATGAGTGAGCTAACTCACATTAATTTGGT 5157
4077 ----- 4076
5158 TGGCTCACTGCGCGCTTTCAGTCCGGAACCTGTGTCGCGAGCTGCAATTAATGATCG 5217
4077 -----TGACCATATGCGGTGTAATAC 4100
5218 GCCAAGCGCGGGAGAGCGGTTTTCGTAATTTGGGCTCTTCGCTTCCTCGCTCACTG 5277
4101 CGCACAGATGCGTAAGGAGAAATACCGCATCAGCGCTCTTCGCTTCCTCGCTCACTG 4160
5278 ACTCGCTGCGCTCGGTGCTGCGTTCGCGTTCGCGGAGCGGTATCAGCTCACTCAAAAGGCGTAA 5337
4161 ACTCGTGGCTCGGTGCTGCGTTCGCGTTCGCGGAGCGGTATCAGCTCACTCAAAAGGCGTAA 4220
5338 TAGCGTTATCCACAGAAATCAGGGGATAACGACAGGAAGAAATGATGAGCAAAAGGCCAGC 5397
4221 TAGCGTTATCCACAGAAATCAGGGGATAACGACAGGAAGAAATGATGAGCAAAAGGCCAGC 4280
5398 AAAAGGCCAGGAACCGTAAAAAGGCGGTTGCTGCGGTTTTTCCATAGGCTCCGCCCCC 5457
4281 AAAAGGCCAGGAACCGTAAAAAGGCGGTTGCTGCGGTTTTTCCATAGGCTCCGCCCCC 4340
5458 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTAT 5517
4341 CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAGGACTAT 4400
5518 AAAGATACAGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGC 5577
4401 AAAGATACAGCGGTTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTGC 4460
5578 CCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCT 5637
4461 CCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCT 4520
5638 CAGGCTGTAGGTATCTCAGTTGCGGTAGGTGTTTCGCTCCAGCTGGGCTGTGTGCAGC 5697
4521 CAGGCTGTAGGTATCTCAGTTGCGGTAGGTGTTTCGCTCCAGCTGGGCTGTGTGCAGC 4580
5698 AACCCCGCTTACGCCCCGACCGTGGCTTAATTCGGTAACTATCGTCTTGAGTCCAAC 5757
4581 AACCCCGCTTACGCCCCGACCGTGGCTTAATTCGGTAACTATCGTCTTGAGTCCAAC 4640
5758 CCGTAAACACGACTTATCGCCACTGGCAGAGCCACTGGTAAAGGATTTAGCAGAGCGA 5817
4641 CCGTAAACACGACTTATCGCCACTGGCAGAGCCACTGGTAAAGGATTTAGCAGAGCGA 4700

QY 5818 GGTATGTAGGCGGTCTACAGAGTTCTTTGAAGTGTGGCTAACTACCGCTACACTAGAA 5877
Db 4701 GGTATGTAGGCGGTCTACAGAGTTCTTTGAAGTGTGGCTAACTACCGCTACACTAGAA 4760
QY 5878 GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTA 5937
Db 4761 GGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTA 4820
QY 5938 GCTCTTGATTCGGGCAAAACAAACCCAGCTGTAGCGGTGGTTTTTTTGTTCGCAAGCAGC 5997
Db 4821 GCTCTTGATTCGGGCAAAACAAACCCAGCTGTAGCGGTGGTTTTTTTGTTCGCAAGCAGC 4880
QY 5998 AGATTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTG 6057
Db 4881 AGATTACCGCGAGAAAAAGGATCTCAAGAGATCTCTTTGATCTTTTCTACGGGTCTG 4940
QY 6058 AGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTTATCAAAAAGA 6117
Db 4941 AGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCATGAGATTTATCAAAAAGA 5000
QY 6118 TCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 6177
Db 5001 TCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 5060
QY 6178 AGTAACTTTGGTCTGACAGTTACCAATCTTAATCAGTGAGGACCTTATCTCAGCGATCT 6237
Db 5061 AGTAACTTTGGTCTGACAGTTACCAATCTTAATCAGTGAGGACCTTATCTCAGCGATCT 5120
QY 6238 GTCTATTTCTGTTTATCCATAGTTGCTTGAATCTCCCGTCTGTGTAGATACTACGATACGG 6297
Db 5121 GTCTATTTCTGTTTATCCATAGTTGCTTGAATCTCCCGTCTGTGTAGATACTACGATACGG 5180
QY 6298 AGGGTCTACATCTGGCCCCAGTGTGCAATGATACCGGGAGACCAAGCTCACCGGCTC 6357
Db 5181 AGGGTCTACATCTGGCCCCAGTGTGCAATGATACCGGGAGACCAAGCTCACCGGCTC 5240
QY 6358 CAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCCGAGCGAGAGTGTCTCTGCAA 6417
Db 5241 CAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCCGAGCGAGAGTGTCTCTGCAA 5300
QY 6418 CTTTATCCGCTTCCATCCAGTCTAATTTGTTGCGGGAAGCTAGAGTAAGTTCG 6477
Db 5301 CTTTATCCGCTTCCATCCAGTCTAATTTGTTGCGGGAAGCTAGAGTAAGTTCG 5360
QY 6478 CAGTTAATAGTTTGGCGAAAGTGTGCAATGCTACAGGC-TCGTGTGTACCGTCTG 6536
Db 5361 CAGTTAATAGTTTGGCGAAAGTGTGCAATGCTACAGGCATCGTGTGTACCGTCTG 5420
QY 6537 CGTTTGGTATGGCTTTCATTCAGCTCCGTTTCCCAACGATCAAGGCGAGTTACATGATCC 6596
Db 5421 CGTTTGGTATGGCTTTCATTCAGCTCCGTTTCCCAACGATCAAGGCGAGTTACATGATCC 5480
QY 6597 CCATGTTTGCAGAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGAGAGTAAGT 6656
Db 5481 CCATGTTTGCAGAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTGAGAGTAAGT 5540
QY 6657 TGGCGCGAGTGTATCAGTATGAGTATGCGAGCATGCAATTAATTTCTTACTGTCTG 6716
Db 5541 TGGCGCGAGTGTATCAGTATGAGTATGCGAGCATGCAATTAATTTCTTACTGTCTG 5600
QY 6717 CATCCGTAAAGTGTCTTTCTGTGACTGTGTAGTACTCAACAAAGTCAATTCAGAGTAAGT 6776
Db 5601 CATCCGTAAAGTGTCTTTCTGTGACTGTGTAGTACTCAACAAAGTCAATTCAGAGTAAGT 5660
QY 6777 GTATCGGCGACCGAGTTGCTCTTCCCGGCGTCAATACGGGATTAATACCGGCCACATA 6836
Db 5661 GTATCGGCGACCGAGTTGCTCTTCCCGGCGTCAACCGGATTAATACCGGCCACATA 5720
QY 6837 CGAGAACTTTAAAGTGTCTCATCATTTGAAACCGTTCTTCGGGGGCAAACTCTCAGGA 6896
Db 5721 CGAGAACTTTAAAGTGTCTCATCATTTGAAACCGTTCTTCGGGGGCAAACTCTCAGGA 5780


```

Qy 6897 TCTTACCGCTCTGAGATCCAGTTCGATGATTAACCCACTCGTGCCACCAACTGATCTTCAG 6956
Db |||||||
Qy 5781 TCTTACCGCTCTGAGATCCAGTTCGATGATTAACCCACTCGTGCCACCAACTGATCTTCAG 5840
Db |||||||
Qy 6957 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 7016
Db |||||||
Qy 5841 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA 5900
Db |||||||
Qy 7017 AAAAGGGAATTAAGGCGGACACGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCAATATT 7076
Db |||||||
Qy 5901 AAAAGGGAATTAAGGCGGACACGGAAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCAATATT 5960
Db |||||||
Qy 7077 ATTGAAGCATTTATCAGGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTAGA 7136
Db |||||||
Qy 5961 ATTGAAGCATTTATCAGGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTTATTAGA 6020
Db |||||||
Qy 7137 AAAATAAACAATAGGGTTTCCGGGCACATTTCCCGAAAAAGTGCACCTGACGCTTAAG 7196
Db |||||||
Qy 6021 AAAATAAACAATAGGGTTTCCGGGCACATTTCCCGAAAAAGTGCACCTGACGCTTAAG 6080
Db |||||||
Qy 7197 AAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCGCTTTTCGTC 7256
Db |||||||
Qy 6081 AAACCATTTATCATGACATTAACCTATAAAAAATAGCGGTATCACGAGGCGCTTTTCGTC 6140
Db |||||||
Qy 7257 T 7257
Db 6141 T 6141

```

RESULT 10

```

US-08-786-531B-3
; Sequence 3, Application US/08786531B
; Patent No. 6541197
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suming
; APPLICANT: Seresina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thonite, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: bgtri
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-786-531B-3

```

```

Query Match 39.9%; Score 3050.2; DB 4; Length 6620;
Best Local Similarity 70.4%; Pred. No. 0;
Matches 4835; Conservative 0; Mismatches 1448; Indels 586; Gaps 33;

```

```

Qy 549 TTTTGAAGAAGCCACCACTGTAGTTTGGCAAGCTAGCTTAAGTAACGCGCATTTTGCAGAGG 608
Db |||||||
Qy 174 TTTTGAAGAAGCCACCACTGTAGTTTGGCAAGCTAGCTTAAGTAACGCGCATTTTGCAGAGG 231
Db |||||||
Qy 609 CATGG-AAAATACATACTAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAC 667
Db |||||||
Qy 232 CATGGAAAATAACATACTAGAAATAGAGAAAGTTTCAGATCAAGGTTAGGAAACAGAGAAAC 291
Db |||||||
Qy 668 AGCAGATATGGGCAACAGGATATCTGTGTAAGCAAGTTCTCTGCCCC-CTCAGAGGCC 726
Db |||||||
Qy 292 AGCTGAATA-----CCAAACAGGATATCTGTGTAAGCGGTTCTCTGCCCCGCTCAGAGGCC 347
Db |||||||
Qy 727 AAGAACAGTTGGAACAGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAAGCAAGTTCTCT 785
Db |||||||
Qy 348 AAGAACAGATGAGACAGCTGAGTGTATGGCCAAACAGGATATCTGTGTAAGCAAGTTCTCT 407
Db |||||||
Qy 786 GCCCGGCTCAGGCGCCCAAGAACAGATGTTCCCGCAGATGCGGTCCCGCTCCAGCAGTTTC 845
Db |||||||
Qy 408 GCCCGGCTCAGGCGCCCAAGAACAGATGTTCCCGCAGATGCGGTCCCGCTCCAGCAGTTTC 467
Db |||||||
Qy 846 TAGAAGAACATCAGATGTTTCCAGGGTCCCAAGGACCTG-AAATGACCTGTGCGCTTA 904
Db |||||||
Qy 468 TAGTGAATCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAATGACCTGTGACCTTA 527
Db |||||||
Qy 905 TTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTCCCGAGCT 964
Db |||||||
Qy 528 TTTGAACCTAACCAATCAGTTTCGCTTCTCGCTTCTGTTGCGCGGCTTCTGCTTCTCGAGCT 587
Db |||||||
Qy 965 CAATAAAGAGGCCACCAACCCCTCACTCGGCGGCCAGTCTCCGATAGACTGGGTGCGCC 1024
Db |||||||
Qy 588 CAATAAAGAGGCCACCAACCCCTCACTCGGCGGCCAGTCTTCCGATAGACTGGGTGCGCC 647
Db |||||||
Qy 1025 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGAGCTCGCTG 1084
Db |||||||
Qy 648 CGGTACCCGTATTTCCCAATAAAGCCTTCTGCTTTGTCATCCGAATCGTGAGCTCGCTG 707
Db |||||||
Qy 1085 ATCCTTGGGAGGCTCTCCTCAGATTTGATTCACCTCCCACTCGGCGGCTTCTTCAATTGGA 1144
Db |||||||
Qy 708 TTTCTTGGGAGGCTCTCCTCTGAGTGAATGACTACCCACGACGGGGGTCTTTTATTGGG 767
Db |||||||
Qy 1145 GGTTCACCGAGATTTTGGAGACCCCTCGCCAGGAGCACCAAGCCCGCCCGCGGAGGTA 1204
Db |||||||
Qy 768 GGCTCGTCCGGATTTGGAGACCCCTCGCCAGGAGCACCAAG-CCACACCGGAGGTA 826
Db |||||||
Qy 1205 AGCTGGCCAGCAACTTATCTGTGTCCTCGATTTGTCTAGTGTCTATGACTGATTTTATG 1264
Db |||||||
Qy 827 AGCTGGCCAGCAACTTATCTGTGTCCTCGATTTGTCTAGTGTCTATGTTTGTATG 886
Db |||||||
Qy 1265 CGCTGCGTCTGCTACTAGTTAGCTTAACCTAGCTCTGATCTGCGGAGCCCGTGTGGAACT 1324
Db |||||||
Qy 887 CGCTGCGTCTGCTACTAGTTAGCTTAACCTAGCTCTGATCTGATCTGCGGAGCCCGTGTGGAACT 946
Db |||||||
Qy 1325 GACGAGTTTCGGAACACCCCGCCCAAGCTTGGGAGAGCTCCAGGAGCTTCCGCGGGCGGT 1384
Db |||||||
Qy 947 GACGAGTTTCGGAACACCCCGCCCAAGCTTGGGAGAGCTCCAGGAGCTTTCGCGGGCGGT 1006
Db |||||||
Qy 1385 TTTTGTGGCCCGGACCTGAGTCTCTAAAAATCCCGATCGTTTTAGGACTCTTTGTGACACCC 1444
Db |||||||
Qy 1007 TTTTGTGGCCCGGACCTGAGGAAGGAGTCTGATGTGGAATCCGACCCCGTCA----- 1057
Db |||||||
Qy 1445 CTTAGAGGAGGATATGTTGTTCTGTTAGAGAGAGAGAACTTAAACAGTTCCCGGCTCC 1504
Db |||||||
Qy 1058 -----GGATATGTTGTTCTGTTAGGAGAGAGAACTTAAACAGTTCCCGGCTCC 1107
Db |||||||

```


Db 3262 GGTCTTGTGATCAGATGATCTGACGAGAGCATCAGG--GCTCGCGCCAGCCGAAC 3319
Qy 3580 CCTCAGGAAGCCTGTCAATGAACCTGAGAAAGATAAGATGCGGAGCCCTACAGTGAAG 3639
Db 3320 TGTTCGCGAGGCTCAAGCGCGCATGCGCGACGGGAGGATCTCGTGTGACCCATGCGG 3379
Qy 3640 ATTGGGATGAAGGCGAGCGCGGAGGGGCAAG--GGCAGATGGCCCTTTACCGAGGTC 3697
Db 3380 ATGCTGTGCTGCCGAATATCATGTGTGGAATATGGCCGCTTTTCTGGATTATCATCGACTGTG 3439
Qy 3698 TCAGTACAGCCACCAAGAGACACTACGACGCCCTTTCATGTCAGGCGCTGCCCTCGCT 3757
Db 3440 GCGGCTGGTGTGCGGACCGCTATCAGACATAGCGTGGCTACCGTATATGCTG 3499
Qy 3758 AA----CTCGACGCGCGCGGATCCGGATTAGTCCAATTTGTTAAAGACAGGATATCAG 3813
Db 3500 AAGAGCTTTGGCGCGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCCGCTCCG 3559
Qy 3814 TGGTCCAGGCTAGTTTGTACTCAACATATACACAGCTGAAGCCTATAGTACGAGC 3873
Db 3560 ATTGACAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTCTTGAGCGGACTCTGGG 3619
Qy 3874 CATGATATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCA 3933
Db 3620 GTTCGATATAAATAAAGATTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCCA 3679
Qy 3934 CCGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCATGG-AAAATACAT 3992
Db 3680 CCGTAGGTTTGGCAAGCTAGCTTAAGTAAAGCCATTTTGAAGGCATGG-AAAATACAT 3739
Qy 3993 AACTGAGATAGAGAGTTTCAGATCAAGGTTAGGAACAGA-GAGACAGCAATATGGGC 4051
Db 3740 AACTGAGATAGAGAGTTTCAGATCAAGGTTAGGAACAGATGGAACAGCTGAATATGGC 3799
Qy 4052 CAACAGCATATCTGTGTAGCAGTCTCTGCCCC-GCTCAGGGCCAGAACAGTTGAA 4110
Db 3800 CAACAGCATATCTGTGTAGCAGTCTCTGCCCGGCTCAGGGCCAGAACAGATGGA 3859
Qy 4111 CAGGAGATATGGGCAACAGGATATCTGTGTAAAGCAGTCTCTGCCCGGCTCAGGGC 4170
Db 3860 CAGCTGATATGGGCAACAGGATATCTGTGTAAAGCAGTCTCTGCCCGGCTCAGGGC 3919
Qy 4171 CAAGAACAGATGTTCCCGAGATGGGTCGCCCTCAGCAGTTCTAGAGAACCATCAGA 4230
Db 3920 CAAGAACAGATGTTCCCGAGATGGGTCGCCCTCAGCAGTTCTAGAGAACCATCAGA 3979
Qy 4231 TGTTCAGGGTGGCCCAAGGACCTGAATGACCTGTGCCTTATTTGAACCTAACCAATC 4290
Db 3980 TGTTCAGGGTGGCCCAAGGACCTGAATGACCTGTGCCTTATTTGAACCTAACCAATC 4039
Qy 4291 AGTTGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCGAGCTCAATAAAGAGCCAC 4350
Db 4040 AGTTGCTTCTCGCTTCTGTTCGCGGCTTCTGTCTCCCGAGCTCAATAAAGAGCCAC 4099
Qy 4351 AACCCCTCACTCGGCGCCAGTCTCCGATAGACTGCGTCCCGGGTACCCCGTGTCT 4410
Db 4100 AACCCCTCACTCGGCGCCAGTCTCCGATGACTGAGTCCGCGGGTACCCCGTGTATC 4159
Qy 4411 CAATAAACCTCTTCAGTTGATGCGGAGGACTCCGTCAGCGGCTGTCTTGGGAGGGTCT- 4469
Db 4160 CAATAAACCTCTTCAGTTGATGCGGAGGACTCCGTCAGCGGCTGTCTTGGGAGGGTCTC 4219
Qy 4470 CTCTGAGTATGACTACCGGTCAGCGGGTCTTTTCAGTTCTCCCACTACAGAGTCT 4529
Db 4220 CTCTGAGTATGACTACCGGTCAGCGGGGCTTTTCATTTTGGGGGCTCGTCCGGGATCG 4279
Qy 4530 CACTAACATTTCTGATGTGCGGAGGAGTCCGTCAGCGCGGTTTTCATTTATAATAAA 4589
Db 4280 GGAGACCCCTGCCAGGAGGACCAACCCAGACCAACCGGGAGGTAGCTGG----- 4328
Qy 4590 TGCAAGAACAGTGTTCCTTCAAGCCAGACTACATCTCAGTCTCGGCTTTATAAAGAA 4649
Db 4329 -----CTGCTCGCGCGTTCGTTGATGA 4352

Qy 4650 TGTGAAAGGGCTCTGTGGAGCTATCTGCCACACGACTTTTTTAAGATTTTTATGCTCTCTGG 4709
Db 4353 CGGTGAAAACTCTGT-ACACATGACAGTCTCCGGAGACGGTCAACAGCTTGTCTGTAAAGCG 4411
Qy 4710 ATGAGGATTTAGTCAATCTATCTCTGCTCTATTTTGTGGCTTCTCCGATTTTAAATTT 4769
Db 4412 ATGCCGGAGCAGACAAGCCCGTCAGGCGCGGTGTTGGCG----- 4459
Qy 4770 CTAGTTTGCACTCCCTCTCTGAGAGCAGCGGATTTGACAGAGTAGTTAATACTCTGAGGCG 4829
Db 4460 ----- 4459
Qy 4830 AGGCTTCTGTAAGAGTTGCTCGGCTCAGTGTGAGATTTTGGCATAAAAAGGGTCT 4889
Db 4460 ----- 4459
Qy 4890 GCCCTGTGTACACACAGATCGGAATCTAGAGTGCATACTCAGAGTCCCCCGGTTCCGG 4949
Db 4460 -----GGTGTGG 4467
Qy 4950 GGCTCTGATCTCAGGCGATCTTTGCTAGAGATCTCTACGCCGAGCAGTCTGGCGCG 5009
Db 4468 GCGCAGCCATGACCCAGTC-----ACGTAGGATAGCGG 4502
Qy 5010 GTACCGAGCTCGAATTCGTAAATCATGTGTCTAGTCTTCTGTGTGAAATTTGTTATCCG 5069
Db 4503 AGTGATATCTGGCTTAACATATGCGGCATCAGACAGATTTGTACTGAGA----- 4550
Qy 5070 CTCAAAATCCACACACATACAGAGCCGGAAGCATAAAGTGTAAAGCCTGGGTGCCTAA 5129
Db 4551 ----- 4550
Qy 5130 TGAGTGAGCTAACTACATTAATTTGCTGTGCTCACTGCCCGCTTTTCCAGTCCGGAAC 5189
Db 4551 ----- 4550
Qy 5190 CTGTCTGCCAGCTGCATTAATGAATCGGCCAACCGCGGGGAGAGCGGTTTGGCTATT 5249
Db 4551 ---GTGCACCATATGCGGTGTGAATACCGCACAGATCGGTAAAGAGAAATACCGCATC 4607
Qy 5250 GGGCGCTCTTCGCTCTCTCGCTCACTGACATCGCTCGGCTCGGCTCGGCTCGGCGGA 5309
Db 4608 AGGCGCTCTTCGCTCTCTCGCTCACTGACCTCGCTCGGCTCGGCTCGGCTCGGCGGA 4667
Qy 5310 GCGGTATCAGCTCACTCAAGCGGTAATACGTTATCCACAGATCAGGGGATACCGCA 5369
Db 4668 GCGGTATCAGCTCACTCAAGCGGTAATACGTTATCCACAGATCAGGGGATACCGCA 4727
Qy 5370 GGAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGAGCCCGCTTG 5429
Db 4728 GGAAGAACATGTGAGCAAAAGGCGCAGCAAAAGGCCAGGAACCGTAAAGAGCCCGCTTG 4787
Qy 5430 CTGCGTTTTTTCATAGCTCCCGCTTGAAGATCAAGGATCAAAAGTCAAGTCAAGT 5489
Db 4788 CTGCGTTTTTTCATAGCTCCCGCTTGAAGATCAAGGATCAAAAGTCAAGTCAAGT 4847
Qy 5490 CAGAGTGGCGAAACCGCAGAGCTATAAGATACAGGCGTTTCCCGCTGGAGCTCC 5549
Db 4848 CAGAGTGGCGAAACCGCAGAGCTATAAGATACAGGCGTTTCCCGCTGGAGCTCC 4907
Qy 5550 CTCTGTGCTCTCTCTGTTCGACCTTACCGGATACCTGTCCGCTTCTCCCT 5609
Db 4908 CTCTGTGCTCTCTCTGTTCGACCTTACCGGATACCTGTCCGCTTCTCCCT 4967
Qy 5610 TCGGAAAGCGTGGCGCTTTCTCATAGCTCAAGTGTAGGTATCTCAGTTCGGTGTAGTGC 5669
Db 4968 TCGGAAAGCGTGGCGCTTTCTCATAGCTCAAGTGTAGGTATCTCAGTTCGGTGTAGTGC 5027
Qy 5670 GTTCGCTCAAGCTGGGCTGTGTGCAAGACCCCGCTTACGCCGAGCGCTGCCTTA 5729
Db 5028 GTTCGCTCAAGCTGGGCTGTGTGCAAGACCCCGCTTACGCCGAGCGCTGCCTTA 5087

Qy	5730	TCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAAGACAGACTTATTCGCCACATGGCAGCA	5789
Db	5088	TCGGTAACTATCGTCTTGTAGTCCAAACCCGGTAAAGACAGACTTATTCGCCACATGGCAGCA	5147
Qy	5790	GCCACTGGTAAACAGGATTTAGCAGAGCGAGTACTGTAGGCGGTGCTACAGAGTCTTCTGAAG	5849
Db	5148	GCCACTGGTAAACAGGATTTAGCAGAGCGAGTACTGTAGGCGGTGCTACAGAGTCTTCTGAAG	5207
Qy	5850	TGTTGGGCTTAACCTACGGCTACACTAGAAAGCAGTATTTGGTATCTCGCTCTGCTCTGAAG	5909
Db	5208	TGTTGGGCTTAACCTACGGCTACACTAGAAAGCAGTATTTGGTATCTCGCTCTGCTCTGAAG	5267
Qy	5910	CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATFCGGCAAAACAAACCAACCGCTGGT	5969
Db	5268	CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATFCGGCAAAACAAACCAACCGCTGGT	5327
Qy	5970	AGCGGTGGTTTTTTTTTGTTCGAAGCAGCAGATTACGGCGCAGAAAAAAGAGATCTCAAGAA	6029
Db	5328	AGCGGTGGTTTTTTTTTGTTCGAAGCAGCAGATTACGGCGCAGAAAAAAGAGATCTCAAGAA	5388
Qy	6030	GATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGG	6089
Db	5388	GATCCTTTGATCTTTTCTACGGGCTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGG	5447
Qy	6090	ATTTTGTGTCATGAGATTATCAAAAGAGATCTTCACTTAGATCCCTTTTAAATATAAAATGA	6149
Db	5448	ATTTTGTGTCATGAGATTATCAAAAGAGATCTTCACTTAGATCCCTTTTAAATATAAAATGA	5507
Qy	6150	AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA	6209
Db	5508	AGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAA	5567
Qy	6210	ATCAGTCAGCAGCACCCTATCTCAGCGATCTGTCATTTTCGTTTCATCCATAGTTGCTGACTC	6269
Db	5568	ATCAGTCAGCAGCACCCTATCTCAGCGATCTGTCATTTTCGTTTCATCCATAGTTGCTGACTC	5627
Qy	6270	CCCGTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTCGCCCCAGTGTGCTGCAATG	6329
Db	5628	CCCGTCGTGTAGATAACTACGATACGGAGGGCTTACCATCTCGCCCCAGTGTGCTGCAATG	5687
Qy	6330	ATACCGCGAGACCCACGCTCACCGGCTCAGAGTTTATCAGCAATAAACACGACGACCGGA	6389
Db	5688	ATACCGCGAGACCCACGCTCACCGGCTCAGAGTTTATCAGCAATAAACACGACGACCGGA	5747
Qy	6390	AGGCGCAGCGCAGAACTGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATGT	6449
Db	5748	AGGCGCAGCGCAGAACTGCTCTGCAACTTTATCCGCCCTCCATCCAGTCTATTAAATGT	5807
Qy	6450	TGCCGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAAATAGTTTGGCACAAGTTGTGCCATT	6509
Db	5808	TGCCGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAAATAGTTTGGCACAAGTTGTGCCATT	5867
Qy	6510	GCTACAGGC--TCGTTGGTGCACGCTCGTCTGTTTGGTATGGTTTCATTTCAGCTCCGGTTC	6568
Db	5868	GCTCAGGCACTCGTGGTGCACGCTCGTCTGTTTGGTATGGTTTCATTTCAGCTCCGGTTC	5927
Qy	6569	CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGTCCCTTC	6628
Db	5928	CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGTTAGTCCCTTC	5987
Qy	6629	GGTCTCCGATCGTTGTTCAGAACTAAGTTGGCCGAGTGTATCACCTCATGTTTATGGCA	6688
Db	5988	GGTCTCCGATCGTTGTTCAGAACTAAGTTGGCCGAGTGTATCACCTCATGTTTATGGCA	6047
Qy	6689	GCACTGCATAAATTCCTTACTGTGCATCGCATCCGTAAGATGCTTTTCTGTGACTGTFGAG	6748
Db	6048	GCACTGCATAAATTCCTTACTGTGCATCGCATCCGTAAGATGCTTTTCTGTGACTGTFGAG	6107
Qy	6749	TACTCAACCAAGTCATTCTCAGATACTGTATGCGGACCGAGTGTCTCTTCCCGCGC	6808
Db	6108	TACTCAACCAAGTCATTCTCAGATACTGTATGCGGACCGAGTGTCTCTTCCCGCGC	6167
Qy	6809	TCAATACGGGATAATACCGGCCCACTAGCAGAACTTTTAAAGTGTCTCATTTTGGAAAA	6868

6168	Db		TCAACGCGGATTAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCTTGGAAAA	6227
6869	Qy	CGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA	6928	
6228	Db	CGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAA	6287	
6929	Qy	CCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTTTCTGGGTGA	6988	
6288	Db	CCCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTACTTTTCCACGCGTTTTCTGGGTGA	6347	
6989	Qy	GCAAAAACAGGAAGGCAGAAATGCGCAAAAAAGGGAATGAAGGCGGCACACGGAATGTTGA	7048	
6348	Db	GCAAAAACAGGAAGGCAGAAATGCGCAAAAAAGGGAATGAAGGCGGCACACGGAATGTTGA	6407	
7049	Qy	ATACTCATACTCTTCTCTTTTCAATATATTGAAGCATTTTATCAGGGTTATTGCTCATG	7108	
6408	Db	ATACTCATACTCTTCTCTTTTCAATATATTGAAGCATTTTATCAGGGTTATTGCTCATG	6467	
7109	Qy	AGCGGATACATATTTGAATGATTTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTT	7168	
6468	Db	AGCGGATACATATTTGATGTATTTAGAAAAATAAAACAAATAGGGGTTCCGCGCACATTT	6527	
7169	Qy	CCCCGAAAAGTGCACCTGACGCTTAAGAAACCATTTATTATCATGACATTTAAACCTATAAA	7228	
6528	Db	CCCCGAAAAGTGCACCTGACGCTTAAGAAACCATTTATTATCATGACATTTAAACCTATAAA	6587	
7229	Qy	AATAGGCGTATCAGAGGCCCTTTTCGTCT	7257	
6588	Db	AATAGGCGTATCAGAGGCCCTTTTCGTCT	6616	

```

RESULT 11
US-08-654-737B-3
; Sequence 3, Application US/08654737B
; Patent No. 6274136
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
; FILE REFERENCE: 13761-703-00 US
; CURRENT APPLICATION NUMBER: US/08/654,737B
; CURRENT FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Retroviral vector derived from Moloney Murine
; OTHER INFORMATION: Leukemia Virus
US-08-654-737B-3

```

	Query Match	35.3%	Score 2703	DB 3	Length 5865
	Best Local Similarity	69.5%	Pred. No. 0		
	Matches 4737	Conservative	0	Mismatches 1065	Indels 1015
	Gaps	38			
Qy	550	TTTGAAGAGCCACCCTGTAGGTTTGCAAGCTAGCTTAAGTAACGCCATTTTGCRAAGC	609		
Db	1	TTTGAAGAGCCACCCTGAGG--TGCAAGCTAGCTTAAGTAACGCCACTTTGCAAGC	58		
Qy	610	ATGG-AAAAATACATACTGAGATAGAGAAAGTTTCAGATCAAGTTTAGGAAACAGAGACA	668		
Db	59	ATGGAAAAATACATACTGAGATAGAAAAAGTTTCAGATCAAGGTCAGGAAACAAGAACA	118		
Qy	669	GCAGATATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCCTGCCCC--GCTCAGGGCCA	727		
Db	119	GCTGAATA-----CAAAACAGGATATCTGTGTAAGCGGTTCTGCCCCGCTCAGGGCCA	174		
Qy	728	AGACAGTTTGGAAACAGGAGAAAT-ATGGGCCAAACAGGATATCTGTGTAAGCAGTTTCCTG	786		

Db 175 AGAAGATGACAGCTGAGTATGGGCGCAACAGGATATCTGGTAAGCAGTTCTCTG 234
Qy 787 CCGCGCTCAGGGCCAGAACAGATGGTCCAGATGCGGTCCCGCTCCAGCAGTTTCT 846
Db 235 CCGCGCTCAGGGCCAGAACAGATGGTCCCGAGATGCGGTCCAGCAGTTTCT 294
Qy 847 AGAGAACCATCAGATGTTTCCAGGTCGCCNAGGACCTG-AAATGACCTGTGCTTAT 905
Db 295 AGTGAATCATCAGATGTTTCCAGGTCGCCNAGGACCTGAAATGACCTGTACTTAT 354
Qy 906 TTGAATTAACCAATCAGTTCGCTCTCTCGTCTCTGTTCCGCGCTCTGCTCCCGAGCTC 965
Db 355 TTGAATTAACCAATCAGTTCGCTCTCTCGTCTCTGTTCCGCGCTCTGCTCCCGAGCTC 414
Qy 966 AATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCCCGATAGACTGCGTCGCC 1025
Db 415 AATAAAGAGCCCAACCCCTCACTCGGCGGCCAGTCTCCCGATAGACTGCGTCGCC 474
Qy 1026 GGGTACCCGTTATCCCAATAAGCCTCTTGTGTTGGATCGGAATCGTGA 1085
Db 475 GGGTACCCGTTATCCCAATAAGCCTCTTGTGTTGGATCGGAATCGTGA 534
Qy 1086 TCCTTGGGAGGTCCTCTCAGATTGATGACTGCCCACTCGGGGTCTTTCATTTGGAG 1145
Db 535 TCCTTGGGAGGTCCTCTCAGATTGATGACTGCCCACTCGGGGTCTTTCATTTGGG 594
Qy 1146 GTTCCACCGAGATTGGAGACCCCTGCGGAGGACCCAGACCCCGCGGAGGTAA 1205
Db 595 GCTCGTCCGGATTGGAGACCCCTGCGGAGGACCCAGACCCCGGAGGTAA 653
Qy 1206 GCTGCCAGCAACTATCTGTCTGTCGCGATGTTCTAGTCTATGACTGATTTATGC 1265
Db 654 GCTGCCAGCAACTATCTGTCTGTCGCGATGTTCTAGTCTATGATTTATGC 713
Qy 1266 GCTGCGTCTGACTAGTACTAGTCTGTTATCTGCGGACCCGCTGGTGA 1325
Db 714 GCTGCGTCTGACTAGTACTAGTCTGTTATCTGCGGACCCGCTGGTGA 773
Qy 1326 ACAGATTCCGAACACCCCGGCGCAACCTCGGAGAGCTGCCAGGACTTCCGGGCGGTT 1385
Db 774 ACAGATTCTGAACACCCCGGCGCAACCTCGGAGAGCTGCCAGGACTTCCGGGCGGTT 833
Qy 1386 TTTGTGGCCCACTGAGTCTCTAAATCCCGATCGTTTAGGACTTTTGTGTCACCCGCC 1445
Db 834 TTTGTGGCCCACTGAGGAGGAGTCTGATGTTGAATCCGACCC----- 878
Qy 1446 TTAGAGGAGGATATGCGTTCTGTAGGAGACAGAACTTAAACAGTTCCCGCTCCG 1505
Db 879 ----CGTCAGGATATGTTCTGTAGGAGACAGAACTTAAACAGTTCCCGCTCCG 934
Qy 1506 TCTGAATTTTGTCTTGGGTTGGACCGAAGCCGCGCGCTCTGT---CTGCTGC 1562
Db 935 TCTGAATTTTGTCTTGGTGGACCGAAGCCGCGCGCTCTGTCTGCTGACGCGTGC 994
Qy 1563 AGCATCGTTCTGTGTGTCTGTCTGACTGTGTTCTGTATTTGTCTGAAATATGGGC 1622
Db 995 AGCATCGTTCTGTGTGTCTGTCTGACTGTGTTCTGTATTTGTCTGAAATATAGGC 1054
Qy 1623 CCGGGCTAGACTGTTACACTCCCTTAAGTTGACCTTGTAGTCTACTGAAAGATGTGAG 1682
Db 1055 C-----AGACTGTTACCACTCCCTTAAGTTGACCTTGTAGTCTACTGAAAGATGTGAG 1108
Qy 1683 CGGATCGCTCACACCACTCGTGTAGTGTCAAGAGAGACGTTGGTTACTCTTCTCTCT 1742
Db 1109 CGGATCGCTCACACCACTCGTGTAGTGTCAAGAGAGACGTTGGTTACTCTTCTCTCT 1168
Qy 1743 GCAGATGGCCAACTTTTAAGCTCGGATGGCGCGGAGACGCTCTTTAAACGAGACCTC 1802
Db 1169 GCAGATGGCCAACTTTTAAGCTCGGATGGCGCGGAGACGCTCTTTAAACGAGACCTC 1228
Qy 1803 ATCACCCAGGTTAAGATCAAGGCTTTTTCACCTGGCCCGCATGGACACCCAGACGAGTC 1862

Db 1229 ATCACCCAGGTTAAGATCAAGGCTTTTTCACCTGGCCCGCATGGACACCCAGACGAGTC 1288
Qy 1863 CCCTACATCGTACCTGGGAAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTTT 1922
Db 1289 CCCTACATCGTACCTGGGAAAGCCTTGGCTTTTGAACCCCTCCCTGGGTCAAGCCCTTT 1348
Qy 1923 GTACACCCCTAAGCCTCGGCTCTCTCTCCATCGGCGCGCTCTCTCCCTTTGAACCT 1982
Db 1349 GTACACCCCTAAGCCTCGGCTCTCTCTCCATCGGCGCGCTCTCTCCCTTTGAACCT 1408
Qy 1983 CCGTCTCGACCCCGCTCGATCTCTCTTATCCAGCCCTCACTCTCTCTTAGGCGCC 2042
Db 1409 CCGTCTCGACCCCGCTCGATCTCTCTTATCCAGCCCTCACTCTCTCTTAGGCGCC 1468
Qy 2043 CCATATGGCCATATGAGATCTTATATGGGCAACCCCGCCCTTTGTAACCTTCCCTGAC 2102
Db 1469 -----GGAAATTCGGCGCGCTACGTAGTCGACTGCG 1499
Qy 2103 CCTGACATGACAGAGTTACTAACAGCCCTCTCTCCAGCTCACTTACAGGCTTCTACT 2162
Db 1500 TGTGGAATGTGTGTCAGTTAGGTTGGAAGTCCCGAGGCTCCCGCAGGC----- 1552
Qy 2163 TAGTCCAGCAGGAGTCTGGAGACCTCTGGCGGCGACCTACCAAGAACAACTGGACCGAC 2222
Db 1553 -----AGAAATGTAACAA----- 1565
Qy 2223 CGGTGTAACCTCACCCCTTACCGAGTCGGCGACACAGTGTGGGTCCGCGACACAGACTA 2282
Db 1566 -----GCATGCACTCAATTAGTCAGCAACAGGTTGTGAAAGTCCCGAGGTTCCCA 1618
Qy 2283 AGAACTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGTCAGCACCCCGCCGCGCC 2342
Db 1619 GCAGGCAAGAGTAGCAAGCATGCATCTCAATTAGTC-----AGCAACCATAGTCCCGC 1673
Qy 2343 TCAAGTAGAGCGCATCGAGCTTGGATACAGCGCGCCAGCTGAGGCTCCGACCCCG 2402
Db 1674 CCCTAACTCCGCCCATCCCGCCCTAACTCGCGCAGTTCGCGCCATTTCTCGGCCCATG 1733
Qy 2403 GGGGTGACCATCTCTAGACTGCGCATGGGATGAGCTGTATCATCTCTTCTTGGTAGC 2462
Db 1734 GCTGACTAATTTTTTTTATTTATGACAGGCGGAGCGCTCGGCTCTGAGCTATTC 1793
Qy 2463 AACAGTACAGGTTTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGAGCTTGAGCGC 2522
Db 1794 AGAAG----TAGTGAGGAGCTTTTGGAGGCTAGGCTTTTGCAAAAAGCTCGAAGAT 1849
Qy 2523 CAGCGTGGGTGACAGAGTGACATCACTGTAAAGCCAGTCAGGATGTGGGTACTTCTGT 2582
Db 1850 CAATTCGATCTGATCAAGA-----GACAGATGAGGATCGTTTCGC 1891
Qy 2583 AGCTTGGTACACAGAGCCAGGTAAAGGCTCCAAAGCTGCTGATCTACTGGACATCCAC 2642
Db 1892 ATGATTAACAAGATGATTTGACAGGAGTTCTCCGCGCGCTTGGGTGGAGAGGCTATTC 1951
Qy 2643 CCGGCACTGCTGTGTCGAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCACCTT 2702
Db 1952 GGCTATGACTGG----- 1963
Qy 2703 CACCATCAGAGCTCTCAGCCAGAGGACATCGCCACCTACTACTGCGCAGCAATAGCCT 2762
Db 1964 ---GCACAAACAGACAACTCGGCTGCTCTGATCGCGCGCTGTTCGGCTGTGAGCGAGGG 2020
Qy 2763 CTATCGGTCTTGGCCCAAGGACCAAGGTGGAATCAACAGGAGTGGCTCAGGATCGGG 2822
Db 2021 CGCCCGGTTCTTTTGTGCAAGCCGACCTGTCCGCTGCTGTAATGAACCTGCAGGACGAG 2080
Qy 2823 TGGATCCGCTCTGGTGGCTCAGGATCGGAGGTCCAACTGCTGGAGAGCGGTGGAGGT 2882
Db 2081 GCAGCGGCTATCTGGCT-----GGCCACGAGCGGCGTTCT 2119
Qy 2883 TGTGCAACCTGGCGGTTCTCGCGCTGTCTCTCGCATCTGGCTTCCGATTTCAACAC 2942
Db 2120 TGGCAGCTGTGCTCGACGTTGTCACTGAAGCGGAAGGAGCTGGCTGCTAT----- 2171

Db 3603 GAGTGATACCTGGCTTAACATATGCGGCAICAGACAGATTTGTACTGAGAGT----- 3653
Qy 5159 GCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTGTCGCCAGCTGCATTAATGAATCGG 5218
Db 3654 -----GCACCATATGCGGTGTGAATACC 3677
Qy 5219 CCAACGCCGGGAGAGCGGTTTGGCTATTTGGGCGCTCTTCGCTTCCCTCGCTCACTGA 5278
Db 3678 GCACAGATGCTAAGGAGAAATACCGCATCAGGCGCTCTTCGCTTCCCTCGCTCACTGA 3737
Qy 5279 CTCGCTGCGCTCGCTGCTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 5338
Db 3738 CTCGCTGCGCTCGCTGCTTCCGCTGCGGAGCGGTATCAGCTCACTCAAAAGCGGTAAT 3797
Qy 5339 ACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCACGA 5398
Db 3798 ACGGTTATCCACAGAAATCAGGGGATAACCGCAGGAAAGAACATGTGAGCAAAAGGCCACGA 3857
Qy 5399 AAAGCCACAGAAACCGTAAAGGCGGCTTCTGCTGCGGTTTTTCATAGGCTCCGCGCCCC 5458
Db 3858 AAAGCCACAGAAACCGTAAAGGCGGCTTCTGCTGCGGTTTTTCATAGGCTCCGCGCCCC 3917
Qy 5459 TGACGAGCATCAAAATACAGCTCAAGTCAGAGGTGGGAAACCCGACAGCACTATA 5518
Db 3918 TGACGAGCATCAAAATACAGCTCAAGTCAGAGGTGGGAAACCCGACAGCACTATA 3977
Qy 5519 AAGATACAGAGGTTTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTTCCGACCCCTGCC 5578
Db 3978 AAGATACAGAGGTTTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTGCTTCCGACCCCTGCC 4037
Qy 5579 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTC 5638
Db 4038 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCATAGCTC 4097
Qy 5639 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGCTGCGCTGTGTGCACGA 5698
Db 4098 ACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGCTGCGCTGTGTGCACGA 4157
Qy 5699 ACCCGCGCTCAGCGCCGACCGCTGCGCTTATCCGTTAACTATCGTGTAGTCCAAACC 5758
Db 4158 ACCCGCGCTCAGCGCCGACCGCTGCGCTTATCCGTTAACTATCGTGTAGTCCAAACC 4217
Qy 5759 GGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTTAAAGGATTAGCAGAGCGAG 5818
Db 4218 GGTAAAGACACGACTTATCGCCACTGGCAGCAGCCACTGTTAAAGGATTAGCAGAGCGAG 4277
Qy 5819 GTATGTAGCGGTGTACAGAGTCTTGAAAGTGGTGGCTTAACTACGCGCTACACTAGAAG 5878
Db 4278 GTATGTAGCGGTGTACAGAGTCTTGAAAGTGGTGGCTTAACTACGCGCTACACTAGAAG 4337
Qy 5879 GACAGTATTTGGTATCTGCGCTGCTGAGCCAGTTTACCTTCGGAAAGAGTTGGTAG 5938
Db 4338 GACAGTATTTGGTATCTGCGCTGCTGAGCCAGTTTACCTTCGGAAAGAGTTGGTAG 4397
Qy 5939 CTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGCA 5998
Db 4398 CTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGTTTTTTTGTTCGAAGCAGCA 4457
Qy 5999 GATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTTTCTACGGGGTCTGA 6058
Db 4458 GATTACGCGCAGAAAAAAGGATCTCAAGAAAGATCTTTTGATCTTTTCTACGGGGTCTGA 4517
Qy 6059 CGCTCAGTGGAAACGAAACTCAGTTAAGGATTTTGTGTCATGAGATTTCAAAAAGGAT 6118
Db 4518 CGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGTGTCATGAGATTTCAAAAAGGAT 4577
Qy 6119 CTTACCTAGATCTTTTAAATTAAGATTAAGTTTTAAATCAATCTAAAGTATATGA 6178
Db 4578 CTTACCTAGATCTTTTAAATTAAGATTAAGTTTTAAATCAATCTAAAGTATATGA 4637
Qy 6179 GTAAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGGGCACTTCTCAGCGATCTG 6238
|||||

Db 4638 GTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGGCACTTATCTCAGCGATCTG 4697
Qy 6239 TCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTGTGTAGATAAATACTACGATACGGGA 6298
Db 4698 TCTATTTTGGTTCATCCATAGTTGCTGACTCCCGTGTGTAGATAAATACTACGATACGGGA 4757
Qy 6299 GGGCTTACCATCTGGCCCGAGTGTCAATGATACCGGAGACCCAGCTCACCGGCTCC 6358
Db 4758 GGGCTTACCATCTGGCCCGAGTGTCAATGATACCGGAGACCCAGCTCACCGGCTCC 4817
Qy 6359 AGATTTATCAGCAATAAACACAGCCAGCGGAAGGCGGAGCGAGAGAGTGTCTCTGCAAC 6418
Db 4818 AGATTTATCAGCAATAAACACAGCCAGCGGAGGCGGAGAGAGTGTCTCTGCAAC 4877
Qy 6419 TTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAAGTACTTCGCC 6478
Db 4878 TTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAAGTACTTCGCC 4937
Qy 6479 AGTTAATAGTTTGGGCAACGTTGTGTGCATTTGTCTACAGGC-TCGTGGTGTACGCTCGTC 6537
Db 4938 AGTTAATAGTTTGGGCAACGTTGTGTGCATTTGTCTGCAAGGATCTGTGGTGTACGCTCGTC 4997
Qy 6538 GTTTGGTATGCTTTCATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCC 6597
Db 4998 GTTTGGTATGCTTTCATTCAGCTCCGCTTCCCAACGATCAAGGCGAGTTACATGATCCCC 5057
Qy 6598 CATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCGATCTGTCAGAGTAAGTT 6657
Db 5058 CATGTTGTGCAAAAAGCGGTTAGTCTCTTCGGTCTCCGATCTGTCAGAGTAAGTT 5117
Qy 6658 GGGCGAGTGTATCACTCATGTTTATGGCAGCATCGATAAATTTCTCTTACTGTCATGCC 6717
Db 5118 GGGCGAGTGTATCACTCATGTTTATGGCAGCATCGATAAATTTCTCTTACTGTCATGCC 5177
Qy 6718 ATCCGTAAGTGTCTTCTGCTGCTGCTGCTCAACCAAGTCATCTGAGTAAGTGTG 6777
Db 5178 ATCCGTAAGTGTCTTCTGCTGCTGCTCAACCAAGTCATCTGAGTAAGTGTG 5237
Qy 6778 TATGCGCGACCGAGTTGCTTCTGCGCGGTCAATACGGGATTAATACCGCGCCACATAG 6837
Db 5238 TATGCGCGACCGAGTTGCTTCTGCGCGGTCAACACGGGATTAATACCGCGCCACATAG 5297
Qy 6838 CAGAACTTTAAAGTGTCTCATTTGGAACGTTTCTCGGGCGCAAACTCTCAAGGAT 6897
Db 5298 CAGAACTTTAAAGTGTCTCATTTGGAACGTTTCTCGGGCGCAAACTCTCAAGGAT 5357
Qy 6898 CTTACCGCTGTGAGATCCAGTTGATTAACCACTGCTGTCACCACTGATCTTCAGC 6957
Db 5358 CTTACCGCTGTGAGATCCAGTTGATTAACCACTGCTGTCACCACTGATCTTCAGC 5417
Qy 6958 ATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCGCAAA 7017
Db 5418 ATCTTTTACTTTTCAACAGCGTTTCTGGGTGAGCAAAACAGGAAGGCAAAATGCGCAAA 5477
Qy 7018 AAAGGGAATTAAGGCGCACACGGAATTTGAATACTCATACTCTTCTCTTTTCAATAATTA 7077
Db 5478 AAAGGGAATTAAGGCGCACACGGAATTTGAATACTCATACTCTTCTCTTTTCAATAATTA 5537
Qy 7078 TTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGAA 7137
Db 5538 TTGAAGCATTTATCAGGGTTATTTGCTCATGAGCGGATACATATTTGAATGTATTTAGAA 5597
Qy 7138 AAATAACAATAGGGGTTCGCGCACATTTTCCCGGAAAGTGCACCTGACGCTCTAGA 7197
Db 5598 AAATAACAATAGGGGTTCGCGCACATTTTCCCGGAAAGTGCACCTGACGCTCTAGA 5657
Qy 7198 AACCATTTATCATGACATTAACTATAAATAAGGCGTATCACGAGGCGCTTTCGCT 7257
Db 5658 AACCATTTATCATGACATTAACTATAAATAAGGCGTATCACGAGGCGCTTTCGCT 5717
Qy 7258 CGCGGTTTCCGTTGATGACCGGTGAAACCTCTGACATGACGCTCCCGGAGACGGTCAAC 7317
Db 5718 TCAAGAAATTCATACAGATCACCGGAAACTGTCTCTCAAAATGTGTCCCGCTCACACTCCC 5777
|||||

Db 1732 AACACAAATCGGCTGCTCTGATCGCGCGTGTTCGGGCTGTACGCGAGGGGCGCCCGG 1791
Qy 2175 -----AAGTCTGGAGACTCTTGGCGGAGCCCTACAAAGAAACAACCTGGACGACCGGTGG 2228
Db 1792 TTCTTTTGTCAAGACCGACCTGTCCGGTGCCTGAATGAATGCTGCAGGACGAGGCGGC 1851
Qy 2229 TAC-----CTCACCTTACCGAGTCCGGGACACAGTGTGGTC 2266
Db 1852 GGTATCGTGTGCGCCAGCGGCGTTCCTTGGCGAGCTGTGCTGACGTTGTCACTG 1911
Qy 2267 CGCGACACACAGACTAAGAACTAGAACTCGCTGGAAGAGACCTTACACAGTCTCTGTG 2326
Db 1912 AAGCGGGAAGGACTGTGCTGTATTGGGGGAAGTCCCGGGCAGAGATCTCTGTATCTC 1971
Qy 2327 ACCACCCCAACCGCCTCAAGTAGACGACATCGCAGCTTGGATACACGCGCCCAACGTG 2386
Db 1972 ACCTTGTCTC-CTGCCGAGAAAGTATCCATCATGCTGATGCAATTCGCGGCGCTGCATAG 2030
Qy 2387 AAGGCTGCGACCCCGGGGTGGACCATCTCTAGACTGCCATGGGATGGAGTGTATCA 2446
Db 2031 CTTGAT-CCGGTACTCTGCCATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAG 2089
Qy 2447 TCCTCTTCTTG-----GTAGCAACAGCTACAGTGTCTCACTCCGACATCCAGCTGACCCA 2501
Db 2090 TACTCGGATGGAGCGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGCT 2149
Qy 2502 GAGCCCAAGCAGCTGAGCGCA-----GCGTGGGTGACAGAGTGACCATCACCT 2551
Db 2150 CGCGCCAGCGAAGCTTTCGCCAGGCTCAAGCGCGCATGCCGACGGGAGGATCTGCT 2209
Qy 2552 GTAAGGCCAGTCAGGATGTGGGTACTTCTGTAGCTTGGTACACAGCAGAACCGAGTAAG 2611
Db 2210 GGTGACCCATGGCGATGCTGCTTCCGAATATCATGTGTGGAATGATGCGCGCTTTCTG 2269
Qy 2612 CTCRAAGCTGTGATCTACTGGACAT--CCACCGGCACACTGTGTGTCACAGCAGATT 2669
Db 2270 ATTCACTACTGTGGCGCGTGGGTGGCGGACGCTATCAGGACATAGCGTTGGGTAC 2329
Qy 2670 CAGCGGTAGCGGTAGCGGTACCGACTTCAACCTTCAACATCAGCAGCCTCCAGCCAGAGGA 2729
Db 2330 CCGTGATATTCTGAAGACTTGGCGGCGAATGGGCTGACCGCTCTCTGTGCTTTAGCG 2389
Qy 2730 CATGCCACCTACTACTGCGCAATATAGCCTCTATCGCTGTTCGGTTCGGCCAAAGGACCAA 2789
Db 2390 TATCGCGCTCCCGATTTCGACGCGATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTG 2449
Qy 2790 GTGGAATC-----AAAGAGGTGGC 2811
Db 2450 AGCGGACTCTGGGTTTGAATGACCGAACCGACCGCCCAACCTGCCATCAGAGAT 2509
Qy 2812 TCAGGATCGGTGGATCGGCTCTCGTCTCAGGATCGAGGTCCAACCTGGTGAGAGC 2871
Db 2510 TTCGATTCACCGCGCCTTCTATGAAGTTGGGCTTCGGAATCGTTTTCGGGACGCC 2569
Qy 2872 GGTGAGGTGTGTCAACCTGGCGGTCCTTGCCTGTCTGTCTGCTGCTGCTGCTGCTGCTG 2931
Db 2570 GGTGATGATCTCTCAGCGGGGATCTCATGCTGGAGTTCTTGCGCCCAACCCCGGCTC 2629
Qy 2932 GATTTACCAATATTGATGATGGTGGGTGAGACAGGACCTTGGAAAGTCTTGAAGTGG 2991
Db 2630 GATCCCTCGAGTTGGTTCA---GCTGTGCTGAGGCTGGACGACCTCGCGAGTT 2685
Qy 2992 ATTGGAGAAATTCACAGATAGCAGTACGATTAACTATGCGCGTCTCTAAAGGATAGA 3051
Db 2686 CTACCGGAGTGAATTCGTTCGGATCCAGGAAACGAGCGGCTATCCGCGATCCA 2745
Qy 3052 TTACAAATATCGGAGACAAACCGCAAGAACCAATTGTTCTCTGCAAAATGGACAGCCTGAGA 3111
Db 2746 TGCCCCCGAAGTGCAGGAGTGGGAGGACGATGCGCGCTTTGGTCGAGGCGGATCTGCT 2805
Qy 3112 CCCGAAGACACCGGGGTCTATTTTGTGCAAGCCTTTTACTCGGCTTCCCTGTGTTGCT 3171
Db 2806 GGTCTTTTCGCGCTCAGAAGCCATAGAGCCACCGCATCCCGCATGCTGCTATTGTC 2865

Qy 3172 TATTGGGGCCAAAGGGAACCCCGGTCAACCGTCTCCAGTGTCTAAAGCCACACAGCGCCAGGG 3231
Db 2866 TTCCCAATCTCTCCCGCTTGTCTCTGCCCCACCCACCCAGAGATAGAAATGACACT 2925
Qy 3232 CGCGACCA-----CCACACCGGGCGCCACCATCGCTCGCAGCCCCCTGTCCCTGC 3283
Db 2926 ACTCAGCAATGCGATGCAATTTCTCATTTATTAGAAAGGACAGTGGAGTGGCACC 2985
Qy 3284 GCCCAGAGGGCGCTCGCGCGCGGGCGCGAGTGCACACAGAGGGGGCTGGACTTCG 3343
Db 2986 TTCCAGGGTCAAGGAAGCGCAGGGGAGGGCAACACAGATGGCTGGCACTAGAGG 3045
Qy 3344 CCCTGGAATCCAAACTCTGTACTCTGTCTGTGATGGAATCTCTTCACTATATGGTGTATTC 3403
Db 3046 CACAGTCAGGCTGATCAGCGAGCTCTAGCATTTAGGTGACACTATAGATAGGGCCCTC 3105
Qy 3404 TCACCTGCTTGTCTCTGAGGTGAAGTTTCAGCAGGAGCGCAGAGCCCCCGGTACACAG 3463
Db 3106 TAGATGATAGCTTAGATCTCGAGCGCGCCACCGGTCCGGATCCGAGCTCGGTAC-CA 3164
Qy 3464 AGGCGCCAGAACCACTCTATAAC---GAGCTCAATCTAGGACGAAGAGAGGATACGATG 3520
Db 3165 AGCTTAAGTTCTCTCTATCACTGATAGGAAATCTCTATCACTGATAGGAGTCTTAT 3224
Qy 3521 TTTTGGCAAGAGACGTGGCGGACCTGTGATGAGTGGGGGAAAGCCGAGAAAGGAAC 3580
Db 3225 ATATCTACTAGTCCGGATCACTATCACTGATAGGAAATCTCTATCACTGATAGGACT 3284
Qy 3581 CTCA--GGAAGGCTGTCAATGAATCTGCAGAAAGATAGATGGGGAGGCTACAGTGA 3638
Db 3285 AGCATGATAGACTCCACCGTACACGCTACCGCCATTTGCGTCAATGGGGCGGAGT 3344
Qy 3639 GATTGGGATGAAGCGCAGCGCGGAGGGCAAGCGCAGTGGCTTTTACAGGGTCT 3698
Db 3345 TGTTACGACATTTTGGAAAGTCCCGTTGATTTTGGTGCAGAAACAACTCCATTTGAGCT 3404
Qy 3699 CAGTACAGCCACCAAGGACACTTACGAGCGCTTTCACATGAGCGCCCTTGCCTTCGCTA 3758
Db 3405 CAATGGGTGGAGACTTGGAAATCCCGTGTAGTCAAACCGCTATCCACGCCCATTTGATGT 3464
Qy 3759 ACTCGACGCGCCCGGATCCGGATTTAGTCAATTTGTTTAAAGACAGATATCAGTGTG 3818
Db 3465 ACTGCCAAAACCGCATCACCATGGTAATAGCATGACTTAATACGTAGTGTACTGCCAAG 3524
Qy 3819 CAGGCTCTAGTTTGTACTCAACAATATCACAGCTGAAGCCTATAGAGTACGAGCCATAG 3878
Db 3525 TAGG--AAAGTCCCATAGGTCTGTGCGCATATGCCAGGGGGCCATTTACCGTCT 3582
Qy 3879 ATAAATAAAAGATTTTATTATTAGTCTCCAGAAAAGGGGGGAATGAAGACCCCACTGT 3938
Db 3583 ATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTTGATGTCTGCCAAGTGGGC 3642
Qy 3939 AGTTTGGCAAGCTAGCTTAAAGTAAAGCCATTTTGAAGGCGATGGAAATACATACTGA 3998
Db 3643 AGTTTACCGTAAATAGTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTAT 3702
Qy 3999 GAATAGAGAGTTCAGATCAGGTTTAGGAAACAGAGAGACAGCAGAAATATGGSCAAACAG 4058
Db 3703 GGGAACTACCTCATTTTGTACGTCAATGGCGGGGGTTCGTTGGCGGTTCAGCCAGGCGG 3762
Qy 4059 GATATCTGTGTAGGACAGTTCC-----TGCCCCGCTCAGGGCCCAAGAACAGTTGGAAACA 4112
Db 3763 GCCATTTACCGTAGTTATGTAAACGCGGAATCCCATATATGGGTATGAACTAATGACCC 3822
Qy 4113 GGAGAAATATGGCCAAACAGGATATCTGTGTGAAGAGTTCCTGCCCCGGGTTCAGGGCCA 4172
Db 3823 CGTAATTGATTACTATTAAATAACTAGTCAATAAATCAATGTCAACGCGTATATCTGGCCG 3882
Qy 4173 AGAACAGATGTTCCCCAGATGCGGTCCCGCTCAGCAGTTTCTTAGA--GAACCATCAGAT 4231
Db 3883 TACATCGCAGAGCAGCGCAAAACCCCTAAACCCCTAAGCAGATTTCTTCATGCAATTTGCGT 3942

QY 4232 GTTTCAGGGTCCCAAGAACCTGAAATGACCTGTGCTTATTTGAACTAACCAATCA 4291
DB 3943 CAAAGCCTTGCTTGTGTAGCTTAAATTTTGTCTCGCGCACTACTCAGCGACCTCCAACAC 4002
QY 4292 GTTTCGCTTCTGCTTCTGTTTGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCA 4351
DB 4003 ACAAGCAGGAGCAGATACCTGGCTTAACATGCGGCATCAGAGCAGATTGTACTGAGAGT 4062
QY 4352 ACCCTCACTCGCGCGCGCAGTCTCCGATA-----GACTCGCTCGCCGGG 4398
DB 4063 CGACCATAGGGATCGGGAGATCCGGGATATAATAAAGATTTTATTTAGTCTCAGAA 4122
QY 4399 TACCCGTGTTCTCAATAAACCCTTTGAGTTGCAATCCGACTCGTGTCTCGCTGTTCT 4458
DB 4123 AAAGGGGGAATGAAGAACCCACACTGTAGTTTGGCAAGCTAGCTTAAAGTAACGCCATT 4182
QY 4459 TGGAGGGTCTCTCTGAGTGAATGACTACCCGTACGGGGTCT-----TTTCA 4506
DB 4183 TTGCAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTTCAGATCAAGGTTCAGGAAC 4242
QY 4507 GTTTCCTCCACCTACACAGTCTCACTAACATTCCTGATGTGCCGAGGACTCCGTGAC 4566
DB 4243 AGATGAACAGCTGAATATGGGCAACACAGGATATCTGTGTAAGCAGATTCTTGCCCCGG 4302
QY 4567 CCGCG-----TTTTGTTTATAATAAATGCAAGAACAGTGTCCCTTCAAGCCA 4616
DB 4303 CTCAGGCCAAGAACAGATGGAACAGCTGAATATGGCCAAACAGATATCTGTGGTAAG 4362
QY 4617 GACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGACTATCTGTC 4676
DB 4363 CAGTTCTCGCCCGCTCAGGGCCAAAGAACAGATGTCGCCAGATGCGGTCAGGCCCTCA 4422
QY 4677 CACACAGCTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATATCTCTCG 4736
DB 4423 GCAGTTTCFAGAGAACCAACAGATGTTCCAGGGTGCCCAAGAGCCTGAATGACCTG 4482
QY 4737 TCTATTTCGTGCTTCTCGCTATTTTAAATTTCTAGTT-----TGCACTCCCTTCTCT 4789
DB 4483 TGCTTATTGAACTAAACCAATCAGTTTCGCTTCTGTTCTGTCGGGCTTCTGCTCC 4542
QY 4790 GAGACCGGGGATTCAGAGTAGTTAACTCTGAGGCGACGGCTTCTGTAAGAAAGTTG 4849
DB 4543 CCGAGCTCAATAAAGAGCCCAACACCTCACTCGGGCGCAGTCTCCGATTTGACTG 4602
QY 4850 CTGGGCTCAGTGTGAGATTTTGGCAATA-----AAGGGTCTG 4890
DB 4603 AGTCGCGCGGTACCCGTGTATCCCAATAAACCTCTTGCACTGTCAGTCTGCTGTC 4662
QY 4891 CCCCTGTGTACAGACAGATCGGAATCTAGAGTGCACTCAGAGTCCCGCGGTTCT---- 4946
DB 4663 TCGCTGTTCTTGGAGGGTCTCTCTGAGTGAATGACTTACCGTTCAGCGGGGCTTTTC 4722
QY 4947 -----CGGGGCTCTGATCTCAGGGGATCTTTGCTTAGAGATCTCTACGCCGAG 4997
DB 4723 ATTTGGGGGCTCGTCCGGATCGGAGACCCCTGCGCCAGGACCCAGACCCACACCG 4782
QY 4998 CATCTGGCGGGTACCGAGCTCGAATTCGTAACTATGTT-CATAGCTGTTTCTGTGTG 5056
DB 4783 GAGGTAACTGTGCTGCTCGCGGTTTTCGGTGTATGACGGTGAACCTCTGACATGCA 4842
QY 5057 AAATTGTATTCGCTCACAAATCCACAAATACAGAGCCGGAAGCAATAAGTGAAGC 5116
DB 4843 GCTCCGGAGACGGTACAGCTTGTCTGTAAAGGATGCGGGAGCAGCAAGCCGGTCA 4902
QY 5117 CTGGGTGCTTAATGAGT-----GAGC 5138
DB 4903 GGGCGGCTCAGCGGGTGTGGGGGTGTCGGGGCGAGCCATGACCCAGTCAAGTAGCA 4962
QY 5139 TAACTACATTAATTTCGTTGGCTCACTGCCGCTTTCAGTCGGGAACCTGTCGTG 5198
DB 4963 TAGCGAGGTGTACTGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCAC 5022
QY 5199 CAGTGCATTAATGAATCGGCCAAACGCGCGGGGAGAGGGGCTTTGCGTATTGGGCGCTCT 5258

DB 5023 CATATCGCGTGTGAATATACCGCACAGATGCGTAGGAGAAAAATACCGATCAGGCGCTCT 5082
QY 5259 TCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTTTCGGCTCGGCGGAGCGGTATCA 5318
DB 5083 TCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTTTCGGCTCGGCGAGCGGTATCA 5142
QY 5319 GCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAAC 5378
DB 5143 GCTCACTCAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAACGACAGAAAGAAC 5202
QY 5379 ATGTGAGCAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAAGCGCGCTTGTGCGCTTT 5438
DB 5203 ATGTGAGCAAAGGCCAGCAAAAGCCAGGAAACCGTAAAAAAGCGCGCTTGTGCGCTTT 5262
QY 5439 TTCCATAGGCTCCGCCGCCCTCGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGG 5498
DB 5263 TTCCATAGGCTCCGCCGCCCTCGACGAGCATCAAAAATTCGACGCTCAAGTCAGAGGTGG 5322
QY 5499 CGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTTGAAGCTCCCTCGTCGCG 5558
DB 5323 CGAAACCCGACAGGACTATAAGATACACAGGCTTTCCCTTGAAGCTCCCTCGTCGCG 5382
QY 5559 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGC 5618
DB 5383 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGC 5442
QY 5619 GTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCC 5678
DB 5443 GTGGCGCTTTCTCATAGCTACGCTGTAGGTATCTCAGTTGCGGTAGGTGCTGCTGCC 5502
QY 5679 AAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGCTAAC 5738
DB 5503 AAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGCTAAC 5562
QY 5739 TATCGCTTTGAGTCCAAACCCGCTTAAGACACAGCTTATCCCACTGCGCAGCAGCCACTG 5798
DB 5563 TATCGCTTTGAGTCCAAACCCGCTTAAGACACAGCTTATCCCACTGCGCAGCAGCCACTG 5622
QY 5799 AACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCT 5858
DB 5623 AACAGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCT 5682
QY 5859 AACTACGCTACACTAGAGGACAGTATTTGCTATCTGCTGCTGCTGAGCAGCTTACC 5918
DB 5683 AACTACGCTACACTAGAGGACAGTATTTGCTATCTGCTGCTGCTGAGCAGCTTACC 5742
QY 5919 TTCCGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGCTGAGCGGTGGT 5978
DB 5743 TTCCGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACCCGCTGCTGAGCGGTGGT 5802
QY 5979 TTTTGTGTTGCAAGCAGAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTG 6038
DB 5803 TTTTGTGTTGCAAGCAGAGATTACGCGCAGAAAAAAGGATCTCAAGAGATCCTTTG 5862
QY 6039 ATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTC 6098
DB 5863 ATCTTTTCTACGGGCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGATTTTGGTC 5922
QY 6099 ATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTTAA 6158
DB 5923 ATGAGATTATCAAAAAGGATCTTCACTAGATCTCTTTTAAATTTAAATTTAAATTTTAA 5982
QY 6159 TCAATCTAAAGTATATATGATGATACTTTGGTCTGACAGTTACCAATGCTTAATCAAGTGA 6218
DB 5983 TCAATCTAAAGTATATATGATGATACTTTGGTCTGACAGTTACCAATGCTTAATCAAGTGA 6042
QY 6219 GCACCTATCTCAGGATCTGCTATTTTCTGTTTCATCATAGTTGCTGACTCCCGCTCGT 6278
DB 6043 GCACCTATCTCAGGATCTGCTATTTTCTGTTTCATCATAGTTGCTGACTCCCGCTCGT 6102
QY 6279 TAGATAACTAGCATACGGGAGGCTTACCATCTGGCCCCAGTGTGCTGAATGATACCGGA 6338

Db 6103 TAGATACTACGATACGGAGGGCTTACCATCTGCCCCAGTGTGCAATGATACCGGA 6162
Qy 6339 GACCACCGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAG 6398
Db 6163 GACCACCGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCGGAGCGGAG 6222
Qy 6399 GCGAAGGTGCTCGCAACTTTATCCCGCTCCATCAGTCTATTAATTTGTCGGGAA 6458
Db 6223 GCGAAGGTGCTCGCAACTTTATCCCGCTCCATCAGTCTATTAATTTGTCGGGAA 6282
Qy 6459 GCTAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTGTTGCCATTTGCTACAGGC 6518
Db 6283 GCTAGAGTAAGTAGTTCGCGAGTTAATAGTTTGGCAACGTTGTTGCCATTTGCTACAGGC 6342
Qy 6519 -TCGTGCTGTCACGCTCGTGTGTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCA 6577
Db 6343 ATCGTGGTGTACGCTCGTGTGTTGGTATGGCTTCATTACGCTCCGGTTCCCAACGATCA 6402
Qy 6578 AGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCGTCCTCG 6637
Db 6403 AGCGAGTTACATGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCGTCCTCG 6462
Qy 6638 ATCGTTGTCAGAGTAAGTTCGCGAGTTTATCAGTCTCATGTTTATGCGAGCACTGCAT 6697
Db 6463 ATCGTTGTCAGAGTAAGTTCGCGAGTTTATCAGTCTCATGTTTATGCGAGCACTGCAT 6522
Qy 6698 AATCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGGTGTACTCAACC 6757
Db 6523 AATCTCTTACTGTCTATGCCATCCGTAAGATGCTTTCTGTGACTGGTGTACTCAACC 6582
Qy 6758 AAGTCATTTCTGAGATAGTGTATGCGGCGACGAGTGTCTCTTGGCCGGGTCTCAATACGG 6817
Db 6583 AAGTCATTTCTGAGATAGTGTATGCGGCGACGAGTGTCTCTTGGCCGGGTCTCAACACGG 6642
Qy 6818 GATATACCGCGCCACATAGCAGACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCG 6877
Db 6643 GATATACCGCGCCACATAGCAGACTTTTAAAGTGTCTCATCTTGGAAACGTTCTTCG 6702
Qy 6878 GGGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGT 6937
Db 6703 GGGGAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGT 6762
Qy 6938 GCACCAACTGATCTTACGATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACA 6997
Db 6763 GCACCAACTGATCTTACGATCTTTTACTTTTACACGCGTTTCTGGGTGAGCAAAAACA 6822
Qy 6998 GGAAGCAAAATGCCGCAAAAAGGAATAGGGCGACACGGAATGTTGAATCTCAT 7057
Db 6823 GGAAGCAAAATGCCGCAAAAAGGAATAGGGCGACACGGAATGTTGAATCTCAT 6882
Qy 7058 CTCCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATAC 7117
Db 6883 CTCCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATGTTCTCATGAGCGGATAC 6942
Qy 7118 ATATTGTAATGTTATGAAAAATAAACAATAGGGGTTCCGCGCACATTTTCCCGGAAA 7177
Db 6943 ATATTGTAATGTTATGAAAAATAAACAATAGGGGTTCCGCGCACATTTTCCCGGAAA 7002
Qy 7178 GTGCCACCTGAGCTGTAGAAACCAATTTATCATGACATTAACCTATAAAAATAGGCGT 7237
Db 7003 GTGCCACCTGAGCTGTAGAAACCAATTTATCATGACATTAACCTATAAAAATAGGCGT 7062
Qy 7238 ATCAGAGCCCTTTCGCT 7257
Db 7063 ATCAGAGCCCTTTCGCT 7082

RESULT 13

US-08-844-274-20/c
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.

; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844.274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence.p3E1.2hs/opd
US-08-844-274-20

Query Match 34.4%; Score 2633.2; DB 3; Length 7560;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 5006 CCGGGTACCGAGCTCGAATTCGTAATCATGGTCTAGCTGTTCTCTGTGAAATTTGTTA 5065
Db 7413 CCGGGTACCGAGCTCGAATTCGTAATCATGGTCTAGCTGTTCTCTGTGAAATTTGTTA 7354
Qy 5066 TCCGCTCACAAATCCACACACATACGCGCGGAGCATAAAGTGAAGCCTCGGGTGC 5125
Db 7353 TCCGCTCACAAATCCACACACATACGCGCGGAGCATAAAGTGAAGCCTCGGGTGC 7294
Qy 5126 CTAATGAGTGAAGTAACTCAATTAATTCGCTTTCGCTCACTGCGCTTTCAGTTCGGG 5185
Db 7293 CTAATGAGTGAAGTAACTCAATTAATTCGCTTTCGCTCACTGCGCTTTCAGTTCGGG 7234
Qy 5186 AAACCTGTGTCGAGCTGCAATTAATGAATCGGCGCAACGCGCGGAGAGCGGTTTCGG 5245
Db 7233 AAACCTGTGTCGAGCTGCAATTAATGAATCGGCGCAACGCGCGGAGAGCGGTTTCGG 7174
Qy 5246 TATTGGGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCG 5305
Db 7173 TATTGGGGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTGCTGCTGCTGCG 7114
Qy 5306 GCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAA 5365
Db 7113 GCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAA 7054
Qy 5366 GCGAGGAAAGAACATGTGAGCAAAAGGCGCAAGAACCCAGAAACCGTAAAAAGGCGCG 5425
Db 7053 GCGAGGAAAGAACATGTGAGCAAAAGGCGCAAGAACCCAGAAACCGTAAAAAGGCGCG 6994
Qy 5426 GTTCTGCGGTTTTTTCATAGGCTCCGCCCCCTTCAGAGATCAGAGATCAGAGATCAGAGCTC 5485
Db 6993 GTTCTGCGGTTTTTTCATAGGCTCCGCCCCCTTCAGAGATCAGAGATCAGAGATCAGAGCTC 6934
Qy 5486 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAG 5545
Db 6933 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAAGCGTTTCCCTTGGAG 6874
Qy 5546 CTCCTCTGCGCTCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCT 5605
Db 6873 CTCCTCTGCGCTCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCCGCTTTCT 6814
Qy 5606 CCTTTCGGAAGCGTGGCGTTTTCTCATAGCTCAAGCTCAGCTGTAGGTATCTCAGTTCGGTGA 5665
Db 6813 CCTTTCGGAAGCGTGGCGTTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGA 6754
Qy 5666 GGTCTGCTCCCAAGCTGGGCTGTGTGCAAGACCCCGCTTTCAGCGCCGACCGCTGCGC 5725
Db 6753 GGTCTGCTCCCAAGCTGGGCTGTGTGCAAGACCCCGCTTTCAGCGCCGACCGCTGCGC 6694
Qy 5726 CTTATCCGGTAACCTATCTGTTGAGTCCAAACCCCGGTAAAGACACGACTTTATCGCCACTGGC 5785

; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-09-598-421-20

Query Match 34.4%; Score 2633.2; DB 4; Length 7560;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 5006 CCGGTACCGAGCTCGAATTCGTAATCATGTCATAGCTGTTCCCTGCTGTGTAATTTGTTA 5065
DB 7413 CCGGTACCGAGCTCGAATTCGTAATCATGTCATAGCTGTTCCCTGCTGTGTAATTTGTTA 7354
QY 5066 TCCGCTCAAAATCCACACACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGGTGC 5125
DB 7353 TCCGCTCAAAATCCACACACATACGAGCCGGAAGCATAAAGTGTAAAGCCTTGGGGTGC 7294
QY 5126 CTAATGAGTGAGTAACCTACATTAATTTGGTGTGCGCTACATGCGCGCTTTCCAGTCGGG 5185
DB 7293 CTAATGAGTGAGTAACCTACATTAATTTGGTGTGCGCTACATGCGCGCTTTCCAGTCGGG 7234
QY 5186 AAACCTGTCGTCAGCTGCATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTCGG 5245
DB 7233 AAACCTGTCGTCAGCTGCATTAATGAATCGGCCAAGCGCGGGAGAGCGGTTTCGG 7174
QY 5246 TATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTCGCTCGCTCGG 5305
DB 7173 TATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGCTCGCTCGG 7114
QY 5306 GCGAGCGGTATCAGTCTCACTAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAA 5365
DB 7113 GCGAGCGGTATCAGTCTCACTAAAGCGGTAAATACGGTTATCCACAGAAATCAGGGGATAA 7054
QY 5366 GCGAGGAAGACATGTCAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGGCGCG 5425
DB 7053 GCGAGGAAGACATGTCAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGGCGCG 6994
QY 5426 GTTGCTGCGGTTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCCACGCTC 5485
DB 6993 GTTGCTGCGGTTTTTCCATAGGCTCCGCCCCCTTGACGAGCATCACAAAATCCACGCTC 6934
QY 5486 AAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGGGGTTTTCCCTCGGAAG 5545
DB 6933 AAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGGGGTTTTCCCTCGGAAG 6874
QY 5546 CTCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCT 5605
DB 6873 CTCCTCGTGGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCT 6814
QY 5606 CCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGA 5665
DB 6813 CCCTTCGGGAAGCGTGGCGCTTTCTCATGTCTCAGCTGTAGGTATCTCAGTTCCGGTGA 6754
QY 5666 GGTGCTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTACGCCCGACCGCTGCGC 5725
DB 6753 GGTGCTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGCTTACGCCCGACCGCTGCGC 6694
QY 5726 CTTATCCCGTAACCTATCTGTTAGTCCAAACCCGCTTAAGACACGACTTATCCGCACTGGC 5785
DB 6693 CTTATCCCGTAACCTATCTGTTAGTCCAAACCCGCTTAAGACACGACTTATCCGCACTGGC 6634
QY 5786 AGCAGCCACTGTATAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTT 5845
DB 6633 AGCAGCCACTGTATAACAGATTTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTT 6574
QY 5846 GAAAGTGGTGGCTTAACCTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCT 5905

DB 6573 GAAGTGGTGGCTTAACCTACGGCTACCTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT 6514
QY 5906 GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCAACCGC 5965
DB 6513 GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAAAACCAACCGC 6454
QY 5966 TGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAAGAAAAAGAGTCTCA 6025
DB 6453 TGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATTACGCGCAAGAAAAAGAGTCTCA 6394
QY 6026 AGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAGTTA 6085
DB 6393 AGAAGATCCTTTGATCTTTCTACGGGTCTGACGCTCAGTGGGAACGAAACTCAGTTA 6334
QY 6086 AGGATTTTGGTCTATGAGATTATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAA 6145
DB 6333 AGGATTTTGGTCTATGAGATTATCAAAAGGATCTTACCTAGATCTCTTTTAAATTTAAA 6274
QY 6146 ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATG 6205
DB 6273 ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTTGGTCTGACAGTTACCAATG 6214
QY 6206 CTTAATCAGTCAGCAGCCTATCTCAGCGATCTGTCTATTTTCTGTTCCATCATAGTTGCGCTG 6265
DB 6213 CTTAATCAGTCAGCAGCCTATCTCAGCGATCTGTCTATTTTCTGTTCCATCATAGTTGCGCTG 6154
QY 6266 ACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTGTC 6325
DB 6153 ACTCCCGTCTGTAGATAACTACGATACGGAGGGCTTACCATCTGCGCCAGTGTGTC 6094
QY 6326 AATGATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAACCCAGCCAGC 6385
DB 6093 AATGATACCGGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAACCCAGCCAGC 6034
QY 6386 CGGAAGGCGCCGAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 6445
DB 6033 CGGAAGGCGCCGAGCGCAGAGTGTCTGCAACTTTTATCCGCTCCATCCAGTCTATTAA 5974
QY 6446 TTGTTGCGGGAAGCTAGAGTAAGTGTCCGCAAGTTAATAGTTTGGCAACGTTGTTGC 6505
DB 5973 TTGTTGCGGGAAGCTAGAGTAAGTGTCCGCAAGTTAATAGTTTGGCAACGTTGTTGC 5914
QY 6506 CATTCGTACAGC-TCGTGCTGTACGCTCGCTGTTGGTATGGCTTCATTCAGTCCGG 6564
DB 5913 CATTCGTACAGC-TCGTGCTGTGTGCTGCTGTTGGTATGGCTTCATTCAGGCGCG 5854
QY 6565 TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTC 6624
DB 5853 TTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTC 5794
QY 6625 CTTCCGCTCTCCGATCGTTGTGCAAGTAAAGTTGGCGCGAGTTATCACATCATGTTAT 6684
DB 5793 CTTCCGCTCTCCGATCGTTGTGCAAGTAAAGTTGGCGCGAGTTATCACATCATGTTAT 5734
QY 6685 GGCAAGCATGCTAATTTCTTACTGTATGCCATCCGCTAAGATGCTTTCTGTGACTGG 6744
DB 5733 GGCAAGCATGCTAATTTCTTACTGTATGCCATCCGTAAGATGCTTTCTGTGACTGG 5674
QY 6745 TGAGTATCAACCAAGTCAATTCGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCC 6804
DB 5673 TGAGTATCAACCAAGTCAATTCGAGAAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCC 5614
QY 6805 GGCGTCAATACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTCTCATCATGTTGG 6864
DB 5613 GGCGTCAATACGGGATTAATACCGCGCACATAGCAGAACTTTAAAGTCTCATCATGTTGG 5554
QY 6865 AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGAT 6924
DB 5553 AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGAT 5494
QY 6925 GTAAACCCACTCGTGACCCCAACTGATCTTTCAGCATCTTTTACTTTTCCAGCGTTTCTGG 6984

```
Db 5493 GTAACCACTGTCACCAACGATCTTTCAGATCTTTTACCTTTTACAGAGGTTTCTGG 5434
Qy 6985 GTGAGCAAAACAGGAAGCAAAATGCCGCAAAAGGAATAAGGCGACAGGAAATG 7044
Db 5433 GTGAGCAAAACAGGAAGCAAAATGCCGCAAAAGGAATAAGGCGACAGGAAATG 5374
Qy 7045 TTGAATACATACACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGCT 7104
Db 5373 TTGAATACATACACTCTTCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGCT 5314
Qy 7105 CATGAGCGATACATATTGAATGATTTAGAAAATAAACAATAAGGGTTCCGGCAC 7164
Db 5313 CATGAGCGATACATATTGAATGATTTAGAAAATAAACAATAAGGGTTCCGGCAC 5254
Qy 7165 ATTTCGCCAAAAGTGCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTA 7224
Db 5253 ATTTCGCCAAAAGTGCCACCTGACGCTCTAAGAAACCAATTATTATCATGACATTAACCTA 5194
Qy 7225 TAAATAAGCGTATCAGAGGCCCTTTCTGTCGCGGCTTTTCGGTGATGACGGTGAATA 7284
Db 5193 TAAATAAGCGTATCAGAGGCCCTTTCTGTCGCGGCTTTTCGGTGATGACGGTGAATA 5134
Qy 7285 CCTCTGACATGACGCTCCGGAGACGTCACAGCTTGTCTGTAAGCGGATGCCGGAG 7344
Db 5133 CCTCTGACATGACGCTCCGGAGACGTCACAGCTTGTCTGTAAGCGGATGCCGGAG 5074
Qy 7345 CAGACAAGCCGTCAGGCGGCTGTCAGCGGCTGTCGCGGCTGTCGCGGCTGCTTAACCTA 7404
Db 5073 CAGACNAGCCGTCAGGCGGCTGTCAGCGGCTGTCGCGGCTGTCGCGGCTGCTTAACCTA 5014
Qy 7405 TCGGCATCAGAGCAGATGTTACTGAGAGTGACCATATGCGGTGTGAAATACCGCACAG 7464
Db 5013 TCGGCATCAGAGCAGATGTTACTGAGAGTGACCATATGCGGTGTGAAATACCGCACAG 4954
Qy 7465 ATGCGTAAGAGAAAATACCGATCAGGCGCATTCGCCATTCAGGCTCGGCACTGTG 7524
Db 4953 ATGCGTAAGAGAAAATACCGATCAGGCGCATTCGCCATTCAGGCTCGGCACTGTG 4894
Qy 7525 GGAAGGGCATCGTTCGCGGCTCTTCGCTATTAGCCAGCTGGCGAAGGGGATGTGC 7584
Db 4893 GGAAGGGCATCGTTCGCGGCTCTTCGCTATTAGCCAGCTGGCGAAGGGGATGTGC 4834
Qy 7585 TGCAAGGCGATTAAGTTGGTTAACGCCAGGTTTCCAGTCACGACGTTGTAACACGAC 7644
Db 4833 TGCAAGGCGATTAAGTTGGTTAACGCCAGGTTTCCAGTCACGACGTTGTAACACGAC 4774
Qy 7645 GGCCAGTGCC 7654
Db 4773 GGCCAGTGCC 4764
```

```
RESULT 15
US-09-205-817A-2
; Sequence 2, Application US/09205817A
; Patent No. 6642028
; GENERAL INFORMATION:
; APPLICANT: Ill, Charles R. et al.
; TITLE OF INVENTION: NOVEL VECTORS AND GENES EXHIBITING
; TITLE OF INVENTION: INCREASED EXPRESSION
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/205,817A
; FILING DATE: 04 DECEMBER 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/067,614
; FILING DATE: 05 DECEMBER 1997
; APPLICATION NUMBER: US 60/071,596
; FILING DATE: 16 JANUARY 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: REMILLARD, JANE E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1006..5376
; US-09-205-817A-2
Query Match 34.3%; Score 2623.4; DB 4; Length 9164;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 5019 TCGAATTCGTATCATGTGTCATAGCTGTTTCTGTGTGAATTTGTATTCGGTCAAAAT 5078
Db 6513 TTGAATTCGTATCATGTGTCATAGCTGTTTCTGTGTGAATTTGTATTCGGTCAAAAT 6572
Qy 5079 CCACACACATACAGAGCGGAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGC 5138
Db 6573 CCACACACATACAGAGCGGAGCATAAAGTGTAAAGCTGGGTGCTTAATGATGAGC 6632
Qy 5139 TAACTCATTAATTTGCGTTCGCTCACTGCCGCTTTCAGTCCGGGAAACCTGCTGTC 5198
Db 6633 TAACTCATTAATTTGCGTTCGCTCACTGCCGCTTTCAGTCCGGGAAACCTGCTGTC 6692
Qy 5199 CAGCTGCAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCT 5258
Db 6693 CAGCTGCAATTAATGAATCGGCAACCGCGGGGAGAGCGGTTTCGCTATTGGGCGCTCT 6752
Qy 5259 TCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTATCA 5318
Db 6753 TCCGCTTCTCGCTCACTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTATCA 6812
Qy 5319 GCTCACTCAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAC 5378
Db 6813 GCTCACTCAAAGCGGTAATACGGTTATCCACAGATCAGGGGATAACGAGGAAGAAC 6872
Qy 5379 ATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGCGCTTGTGCGGTTT 5438
Db 6873 ATGTGAGCAAAAGGCGCAGCAAAAGCCAGGAAACCGTAAAGAGCGCGCTTGTGCGGTTT 6932
Qy 5439 TTCCATAGGCTCCGCCCCCTCGAGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGG 5498
Db 6933 TTCCATAGGCTCCGCCCCCTCGAGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGG 6992
Qy 5499 CGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCTCGGAAGCTCCCTCGTGGC 5558
Db 6993 CGAAACCCGACAGGACTATAAGATACAGGGGTTTCCCTCGGAAGCTCCCTCGTGGC 7052
Qy 5559 TCTCTGTTTCGACCCCTCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGCGAAGC 5618
Db 7053 TCTCTGTTTCGACCCCTCGCTTACCGATACCTGTCGCGCTTCTCCCTTCGCGAAGC 7112
Qy 5619 GTGGCGCTTCTCATAGCTCAAGTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCC 5678
```


Db 7113 GTGGCGCTTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTCGGTGTAGGTGGTTCGCTCC 7172
Qy 5679 AAGTGGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAAAC 5738
Db 7173 AAGTGGGCTGTGTGCAAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAAAC 7232
Qy 5739 TATCGTCTTGTAGTCCAAACCCGGTAAAGACAGCACTTATCGCCACTGGCAGAGCCACTGGT 5798
Db 7233 TATCGTCTTGTAGTCCAAACCCGGTAAAGACAGCACTTATCGCCACTGGCAGAGCCACTGGT 7292
Qy 5799 AACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCT 5858
Db 7293 AACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCT 7352
Qy 5859 AACTACGGCTACCTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTACC 5918
Db 7353 AACTACGGCTACCTAGAGGACAGTATTTGGTATCTCGCTCTGCTGAAAGCCAGTTACC 7412
Qy 5919 TTGGAAAAAGATTGGTAGTCTCTTGTATCCGGCAAAACAAACCGCTGGTAGCGGTGGT 5978
Db 7413 TTGGAAAAAGATTGGTAGTCTCTTGTATCCGGCAAAACAAACCGCTGGTAGCGGTGGT 7472
Qy 5979 TTTTGTGTTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCCCTTG 6038
Db 7473 TTTTGTGTTGCAAGCAGCAGATTACCGCGCAGAAAAAAGGATCTCAAGAGATCCCTTG 7532
Qy 6039 ATCTTTTCTACGGGCTCTGACGCTCAGTGGGAACGAAAACTCACTGTTAAGGATTTGGTC 6098
Db 7533 ATCTTTTCTACGGGCTCTGACGCTCAGTGGGAACGAAAACTCACTGTTAAGGATTTGGTC 7592
Qy 6099 ATGAGATTATCAAAAAAGGATCTTACCTAGATTCCTTTTAAATTAATAAATGAAGTTTAAA 6158
Db 7593 ATGAGATTATCAAAAAAGGATCTTACCTAGATTCCTTTTAAATTAATAAATGAAGTTTAAA 7652
Qy 6159 TCAATCTAAAGTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAG 6218
Db 7653 TCAATCTAAAGTATATAGTAACTTGGTCTGACAGTTACCAATGCTTAAATCAGTGAG 7712
Qy 6219 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCATAGTTGCTGACTCCCGTCTGT 6278
Db 7713 GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCACTCATAGTTGCTGACTCCCGTCTGT 7772
Qy 6279 TAGATACTACGATACGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 6338
Db 7773 TAGATACTACGATACGAGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGGA 7832
Qy 6339 GACCCAGCTCAGCGCTCCAGATTATACAGATTAACAGACGACCGGGAAGGCGGAG 6398
Db 7833 GACCCAGCTCAGCGCTCCAGATTATACAGATTAACAGACGACCGGGAAGGCGGAG 7892
Qy 6399 CGCAGAGTGGTCTTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGA 6458
Db 7893 CGCAGAGTGGTCTTGCAACTTTATCCGCTCCATCCAGTCTATTAAATTGTTGCGGGA 7952
Qy 6459 GCTAGAGTAAGTGTGCGCAGTTAAATAGTTTGGCAGAACGTTGTTGCCATTGCTACAGG 6518
Db 7953 GCTAGAGTAAGTGTGCGCAGTTAAATAGTTTGGCAGAACGTTGTTGCCATTGCTACAGG 8012
Qy 6519 -TCGTGTGTCAGCGTCTGTTGGTAGGCTTCACTTACGCTCGGTTCCCAACGATCA 6577
Db 8013 ATCGTGTGTACGCTCGTCTGTTGGTAGGCTTCACTTACGCTCGGTTCCCAACGATCA 8072
Qy 6578 AGGCGAGTTACATGATCCCGCATGTTGCAAAAAAGCGTTAGCTCCTTCGTTCCCTCG 6637
Db 8073 AGGCGAGTTACATGATCCCGCATGTTGCAAAAAAGCGTTAGCTCCTTCGTTCCCTCG 8132
Qy 6638 ATCGTGTGTCAGAGTAAGTGTGCGCGCAGTTTATCACTCATGTTTATGGCAGCACTGCAT 6697
Db 8133 ATCGTGTGTCAGAGTAAGTGTGCGCGCAGTTTATCACTCATGTTTATGGCAGCACTGCAT 8192
Qy 6698 AATTCTCTTATGTCATGCCATCCGTAAGATGCTTTTCTGTCAGTGTGAGTACTCAACC 6757
Db 8193 AATTCTCTTATGTCATGCCATCCGTAAGATGCTTTTCTGTCAGTGTGAGTACTCAACC 8252

Qy 6758 AAGTCATTTCTGAGAAATAGTGTATCGCGCAGACCGAGTTGCTCTTGCCTCCGCGTCAATACGG 6817
Db 8253 AAGTCATTTCTGAGAAATAGTGTATCGCGCAGACCGAGTTGCTCTTGCCTCCGCGTCAATACGG 8312
Qy 6818 GATTAATCCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAGTCTTCGG 6877
Db 8313 GATTAATCCGCGCCACATAGCAGAACTTTAAAGTGTCTCATCATTTGGAAGAGTCTTCGG 8372
Qy 6878 GGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCACTCGT 6937
Db 8373 GGGCGAAAACTCTCAAGGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCCACTCGT 8432
Qy 6938 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTCCTGGGTGAGCAAAAAACA 6997
Db 8433 GCACCCAACTGATCTTTCAGCATCTTTTACTTTTCCACAGCGTTCCTGGGTGAGCAAAAAACA 8492
Qy 6998 GGAAGGCAAAATGCGCAAAAAAGGAAATAGGCGGACACGGAATGTTGAATACTCAT 7057
Db 8493 GGAAGGCAAAATGCGCAAAAAAGGAAATAGGCGGACACGGAATGTTGAATACTCAT 8552
Qy 7058 CTCTTCTTTTCAATATTTTGAAGCATTTTATCAGGGTATTTGCTCATGAGCGGATAC 7117
Db 8553 CTCTTCTTTTCAATATTTTGAAGCATTTTATCAGGGTATTTGCTCATGAGCGGATAC 8612
Qy 7118 ATATTTGAATGTTTATAGAAAAATAAACAATAGGGTTCGCGCACATTTTCCCGGAAAA 7177
Db 8613 ATATTTGAATGTTTATAGAAAAATAAACAATAGGGTTCGCGCACATTTTCCCGGAAAA 8672
Qy 7178 GTGCCACTGACGCTTAAAGAAACCATTTATTCATGATTAACCTATAAATAAGGCT 7237
Db 8673 GTGCCACTGACGCTTAAAGAAACCATTTATTCATGATTAACCTATAAATAAGGCT 8732
Qy 7238 ATCAGAGGCGCTTTCGTCGCGCTTCGTCGATGACCGTGAACCTCTGACACATG 7297
Db 8733 ATCAGAGGCGCTTTCGTCGCGCTTCGTCGATGACCGTGAACCTCTGACACATG 8792
Qy 7298 CAGTCCCGGAGACGCTCAGACGTTGTCTGTAAAGCGATGCGGGAGCAGACAGCCCGT 7357
Db 8793 CAGTCCCGGAGACGCTCAGACGTTGTCTGTAAAGCGATGCGGGAGCAGACAGCCCGT 8852
Qy 7358 CAGGCGCGTCAAGCGGCTGTTGGCGGTGTCGCGGCTGGCTTAACTATGCGGCTACAG 7417
Db 8853 CAGGCGCGTCAAGCGGCTGTTGGCGGTGTCGCGGCTGGCTTAACTATGCGGCTACAG 8912
Qy 7418 CAGATTCTGAGAGTGCACCATATCGGTGTCAATACCGCACAGATGCTGAAGGAGA 7477
Db 8913 CAGATTCTGAGAGTGCACCATATCGGTGTCAATACCGCACAGATGCTGAAGGAGA 8972
Qy 7478 AAATACCGCATCAGCGCCCATTTCCGCTTTCAGGCTGCGCAACTGTTGGAAAGGCGATCG 7537
Db 8973 AAATACCGCATCAGCGCCCATTTCCGCTTTCAGGCTGCGCAACTGTTGGAAAGGCGATCG 9032
Qy 7538 GTGCGGCGCTTTCGCTTATTAACCGCAGTGGCGAAAGGGGATGTTGTCGAAGCGATTA 7597
Db 9033 GTGCGGCGCTTTCGCTTATTAACCGCAGTGGCGAAAGGGGATGTTGTCGAAGCGATTA 9092
Qy 7598 AGTTGGTAAACCGCAGGCTTTTCCAGTCAGGATGTTTAAACGACCGGCGAGTGC 7654
Db 9093 AGTTGGTAAACCGCAGGCTTTTCCAGTCAGGATGTTTAAACGACCGGCGAGTGC 9149

Search completed: May 25, 2005, 19:58:51
Job time : 1017 secs

THIS PAGE BLANK (USPTO)

QY 61 GCAGCCGAGACAGTGAAGAAACCGTTAAACCGTTGTTTAAATAAACTGAATATTTT 120
 DB |||||
 QY 61 GCAGCCGAGACAGTGAAGAAACCGTTAAACCGTTGTTTAAATAAACTGAATATTTT 120
 DB |||||
 QY 121 AGAGTCATTTCTTTTGTGAGAAAGTACATTTGGCAGCTGAAGAGGCCAAAGCAATCTGTG 180
 DB |||||
 QY 121 AGAGTCATTTCTTTTGTGAGAAAGTACATTTGGCAGCTGAAGAGGCCAAAGCAATCTGTG 180
 DB |||||
 QY 181 GAAAGCCGAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCTGGCGTGTACCTTAAGGGTTT 240
 DB |||||
 QY 181 GAAAGCCGAGGCTGGAGGCCAGCAGTTTGCATCCCTCTCTGGCGTGTACCTTAAGGGTTT 240
 DB |||||
 QY 241 CTTAATTTGTGTGTTTCTTAAATCTTCCAGAGGTTTGTCTCATTTCACTTCCACTTCGGTG 300
 DB |||||
 QY 241 CTTAATTTGTGTGTTTCTTAAATCTTCCAGAGGTTTGTCTCATTTCACTTCCACTTCGGTG 300
 DB |||||
 QY 301 CACAATPACTTGGACGGGATTTATCTGTCTTAGCATCTATCGTGGGCCCTTCGATTTAGGCG 360
 DB |||||
 QY 301 CACAATPACTTGGACGGGATTTACTGTCTTAGCATCTATCGTGGGCCCTTCGATTTAGGCG 360
 DB |||||
 QY 361 TGAACCTGAGGCCACTTCTTACGCTTGTAGGAGGACACAAGCAGCAGAGAGGCTGA 420
 DB |||||
 QY 361 TGAACCTGAGGCCACTTCTTACGCTTGTAGGAGGACACAAGCAGCAGAGAGGCTGA 420
 DB |||||
 QY 421 CCGCGCAGACCTGTGGGCATTTTAAACAGGGCTCTCTGGTCTGTGGGAGGCGAGGCTTA 480
 DB |||||
 QY 421 CCGCGCAGACCTGTGGGCATTTTAAACAGGGCTCTCTGGTCTGTGGGAGGCGAGGCTTA 480
 DB |||||
 QY 481 CATAGGTCGAAATTTAGAAATATAAATAAAGCCCATATCAATTTGTCTATCTTTTATTA 540
 DB |||||
 QY 481 CATAGGTCGAAATTTAGAAATATAAATAAAGCCCATATCAATTTGTCTATCTTTTATTA 540
 DB |||||
 QY 541 AGCTCAAGTTTGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATT 600
 DB |||||
 QY 541 AGCTCAAGTTTGAAGAGCCCACTGTAGGTTTGGCAAGCTAGCTTAAAGTAAAGCCATT 600
 DB |||||
 QY 601 TTGCAAGGATGAGAAATACATAACTGAGAAATAGAGAAAGTTCAGATCAAGAGTTTAGGAACA 660
 DB |||||
 QY 601 TTGCAAGGATGAGAAATACATAACTGAGAAATAGAGAAAGTTCAGATCAAGAGTTTAGGAACA 660
 DB |||||
 QY 661 GAGAGACAGAGATATGGGCCAAACACAGGATATCTGTGGTAAGCAGTTCTCTGCCCGCTC 720
 DB |||||
 QY 661 GAGAGACAGAGATATGGGCCAAACACAGGATATCTGTGGTAAGCAGTTCTCTGCCCGCTC 720
 DB |||||
 QY 721 AGGGCCCAAGAACAGTTTGGAAACAGGAGAAATAGGGCCAAACAGGATATCTGTGGTAAGCAG 780
 DB |||||
 QY 721 AGGGCCCAAGAACAGTTTGGAAACAGGAGAAATAGGGCCAAACAGGATATCTGTGGTAAGCAG 780
 DB |||||
 QY 781 TTCTCTGCGCGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCA 840
 DB |||||
 QY 781 TTCTCTGCGCGCTCAGGGCCCAAGAACAGATGTCCTCCAGATGCGGTCCCGCCCTCAGCA 840
 DB |||||
 QY 841 GTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGAGGACCTGAAATGACCCCTGTGC 900
 DB |||||
 QY 841 GTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCAAGAGGACCTGAAATGACCCCTGTGC 900
 DB |||||
 QY 901 CTTAATTTGAACCTAAACCAATCAGTTCTGCTTCTGCTTCTGTTTGGCGGCTTCTGCTCCCG 960
 DB |||||
 QY 901 CTTAATTTGAACCTAAACCAATCAGTTCTGCTTCTGCTTCTGTTTGGCGGCTTCTGCTCCCG 960
 DB |||||
 QY 961 AGCTCAATTAAGAGCCCAACCCCTCACTGGCGGCGCAGTCTCCGATAGACTGCGT 1020
 DB |||||
 QY 961 AGCTCAATTAAGAGCCCAACCCCTCACTGGCGGCGCAGTCTCCGATAGACTGCGT 1020
 DB |||||
 QY 1021 CGCCCGGCTACCGGATTTCCCAATAAAGGCTCTTGTGTTTGCATCCGAATCGTGACTC 1080
 DB |||||
 QY 1021 CGCCCGGCTACCGGATTTCCCAATAAAGGCTCTTGTGTTTGCATCCGAATCGTGACTC 1080
 DB |||||
 QY 1081 GCTGATCTTTGGAGGGTCTCTCAGATTGATTTGACTGCCACCTCGGGGGTCTTTCAATT 1140
 DB |||||
 QY 1081 GCTGATCTTTGGAGGGTCTCTCAGATTGATTTGACTGCCACCTCGGGGGTCTTTCAATT 1140
 DB |||||

QY 1141 TGGAGGTTCCACCGAGATTTTGGAGACCCCTGTCCCGAGGACCAACGACCCCGCCGCGGA 1200
 DB |||||
 QY 1141 TGGAGGTTCCACCGAGATTTTGGAGACCCCTGTCCCGAGGACCAACGACCCCGCCGCGGA 1200
 DB |||||
 QY 1201 GGTAAAGCTGGCCAGCAACTTATCTGTGTCTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTT 1260
 DB |||||
 QY 1201 GGTAAAGCTGGCCAGCAACTTATCTGTGTCTGTCTGTCCGATTGTCTAGTGTCTATGACTGATTT 1260
 DB |||||
 QY 1261 TATGGCGCTCGGTCTAGTACTAGTAACTAGTCTGTATCTGTGCGGACCCGTTGGTG 1320
 DB |||||
 QY 1261 TATGGCGCTCGGTCTAGTACTAGTAACTAGTCTGTATCTGTGCGGACCCGTTGGTG 1320
 DB |||||
 QY 1321 AACTGACGAGTTTCGAAACACCCGCGCGCAACCCCTGGGAGACGCTCCAGGGAATTCGGGG 1380
 DB |||||
 QY 1321 AACTGACGAGTTTCGAAACACCCGCGCGCAACCCCTGGGAGACGCTCCAGGGAATTCGGGG 1380
 DB |||||
 QY 1381 CCGTTTTTGTGGCCCGGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTTGGTGAC 1440
 DB |||||
 QY 1381 CCGTTTTTGTGGCCCGGACCTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTTGGTGAC 1440
 DB |||||
 QY 1441 CCCCCTTAGAGGAGGATATGTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTTCCCG 1500
 DB |||||
 QY 1441 CCCCCTTAGAGGAGGATATGTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTTCCCG 1500
 DB |||||
 QY 1501 CTCCGCTCTGAATTTTGTGTTTTCGGTTTGGGACCCGAGCGCGCGCTCTTGTCTGCT 1560
 DB |||||
 QY 1501 CTCCGCTCTGAATTTTGTGTTTTCGGTTTGGGACCCGAGCGCGCGCTCTTGTCTGCT 1560
 DB |||||
 QY 1561 GCAGCATCGTCTGTGTTGTCTCTGTCTGACTGTCTTCTGTATTTGTCTGAAAAATATGG 1620
 DB |||||
 QY 1561 GCAGCATCGTCTGTGTTGTCTCTGTCTGACTGTCTTCTGTATTTGTCTGAAAAATATGG 1620
 DB |||||
 QY 1621 GCCCGGCTTAGACTCTTACCACTCCCTTAAAGTTTACCTTACCTTACCTTACCTTACCTTACCT 1680
 DB |||||
 QY 1621 GCCCGGCTTAGACTCTTACCACTCCCTTAAAGTTTACCTTACCTTACCTTACCTTACCTTACCT 1680
 DB |||||
 QY 1681 AGCGATCGCTCACCAACAGTCTGTAGTGTCTCAAGAGAGAGAGTGTGGGTACCTTCTGCT 1740
 DB |||||
 QY 1681 AGCGATCGCTCACCAACAGTCTGTAGTGTCTCAAGAGAGAGAGTGTGGGTACCTTCTGCT 1740
 DB |||||
 QY 1741 CTGCAAGATTTGGCAACCTTTTAACTGCGATGTCGCGAGAGAGAGACCTTTTAAACGAGACC 1800
 DB |||||
 QY 1741 CTGCAAGATTTGGCAACCTTTTAACTGCGATGTCGCGAGAGAGAGACCTTTTAAACGAGACC 1800
 DB |||||
 QY 1801 TCATCACCAGGTTAAGATCAAGTCTTTTACCTGCGCGCATCGACACCCAGACCCAGG 1860
 DB |||||
 QY 1801 TCATCACCAGGTTAAGATCAAGTCTTTTACCTGCGCGCATCGACACCCAGACCCAGG 1860
 DB |||||
 QY 1861 TCCCTTACATCTGACCTGGGAAGCTTTGGCTTTTGAACCCCTCTCTGGGTCAAGCCCT 1920
 DB |||||
 QY 1861 TCCCTTACATCTGACCTGGGAAGCTTTGGCTTTTGAACCCCTCTCTGGGTCAAGCCCT 1920
 DB |||||
 QY 1921 TTGTACACCTTAAGCCTCGGCTCTCTTCTTCTTCAATTCGCGCCGCTCTCTCCGCTTTGAAC 1980
 DB |||||
 QY 1921 TTGTACACCTTAAGCCTCGGCTCTCTTCTTCTTCAATTCGCGCCGCTCTCTCCGCTTTGAAC 1980
 DB |||||
 QY 1981 CTCTCTGTCACCCGCGCTCGATCTCTCTTATCCAGCCCTCTCTCTCTCTAGGG 2040
 DB |||||
 QY 1981 CTCTCTGTCACCCGCGCTCGATCTCTCTTATCCAGCCCTCTCTCTCTCTAGGG 2040
 DB |||||
 QY 2041 CCCCCTATGCGCATATGAGATCTTATATGGGCAACCCCGCCCTCTTAAACTTCCCTCTG 2100
 DB |||||
 QY 2041 CCCCCTATGCGCATATGAGATCTTATATGGGCAACCCCGCCCTCTTAAACTTCCCTCTG 2100
 DB |||||
 QY 2101 ACCCTGACATGAAGAAGTTACTAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160
 DB |||||
 QY 2101 ACCCTGACATGAAGAAGTTACTAACAGCCCTCTCTCTCAAGCTCACTTACAGGCTTCTA 2160
 DB |||||
 QY 2161 CTTAGTCAGCAGAGTCTGGAGACCTTGGCGGAGCCTTACAGACACTTGGACCG 2220
 DB |||||
 QY 2161 CTTAGTCAGCAGAGTCTGGAGACCTTGGCGGAGCCTTACAGACACTTGGACCG 2220
 DB |||||
 QY 2221 ACCGGTGGTACCTCACCCCTTACCAGTGGCGGACACAGTGTGGGTCCGCGCACACAGAC 2280
 DB |||||

Db 2221 ACCGGTGTACTTACCCCTTACCGAGTCGGGACACAGTGTGGGTCCGGCCGACACAGAC 2280
Qy 2281 TAAGAACCTTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGTGACACACCCGACCCG 2340
Db 2281 TAAGAACCTTAGAACCTCGCTGGAAGGACCTTACACAGTCTCTGTGTGACACACCCGACCCG 2340
Qy 2341 CCTCAAAGTAGAGCGGATCGAGCTTGGATACACCGCCGCCACGCGTGAAGGCTGCGGACCC 2400
Db 2341 CCTCAAAGTAGAGCGGATCGAGCTTGGATACACCGCCGCCACGCGTGAAGGCTGCGGACCC 2400
Qy 2401 CGGGGTGACCATCTCTAGACTGCGCATGGATGGAGTGTATCATCTCTTCTTGGTA 2460
Db 2401 CGGGGTGACCATCTCTAGACTGCGCATGGATGGAGTGTATCATCTCTTCTTGGTA 2460
Qy 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGAGC 2520
Db 2461 GCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGAGC 2520
Qy 2521 GCCAGCGTGGTGAACAGAGTGAACATCACTGTGAAGCCAGTCAAGATGTGGGTACTTCT 2580
Db 2521 GCCAGCGTGGTGAACAGAGTGAACATCACTGTGAAGCCAGTCAAGATGTGGGTACTTCT 2580
Qy 2581 GTAGCTTGGTACCAGCAGAGCCAGGTAAAGCTCCAAAGCTGCTGATCTACTGGACATCC 2640
Db 2581 GTAGCTTGGTACCAGCAGAGCCAGGTAAAGCTCCAAAGCTGCTGATCTACTGGACATCC 2640
Qy 2641 ACCGGCACACTGGTGTGCAAGCAGATTACGCGGTAGCGGTAGCGGTACCGACTTACC 2700
Db 2641 ACCGGCACACTGGTGTGCAAGCAGATTACGCGGTAGCGGTAGCGGTACCGACTTACC 2700
Qy 2701 TTCAACATCAGCAGCTTCCAGCCAGAGGACATCGCCACTACTACTGCGCAGCAATATAGC 2760
Db 2701 TTCAACATCAGCAGCTTCCAGCCAGAGGACATCGCCACTACTACTGCGCAGCAATATAGC 2760
Qy 2761 CTCTATCGGTGTTCCGGCAGAGGACCAAGGTGGAATCAAACGAGGTGGCTCAGGATCG 2820
Db 2761 CTCTATCGGTGTTCCGGCAGAGGACCAAGGTGGAATCAAACGAGGTGGCTCAGGATCG 2820
Qy 2821 GGTGATCCGGCTCTGGTGGCTCAGGATCGGAGGTCCAACTGGTGGAGCGGTGGAGGT 2880
Db 2821 GGTGATCCGGCTCTGGTGGCTCAGGATCGGAGGTCCAACTGGTGGAGCGGTGGAGGT 2880
Qy 2881 GTTGTCAACCTGGCGGCTCTGGCTGCTGCTCGCATCTGGCTTCGATTTTACC 2940
Db 2881 GTTGTCAACCTGGCGGCTCTGGCTGCTGCTCGCATCTGGCTTCGATTTTACC 2940
Qy 2941 ACATATTGGATGAGTTGGGTGAGACAGGCACTTGAAAAGTCTTGAAGTGAATTGGAGAA 3000
Db 2941 ACATATTGGATGAGTTGGGTGAGACAGGCACTTGAAAAGTCTTGAAGTGAATTGGAGAA 3000
Qy 3001 ATTCAATCCAGATAGCAGTACGATTAACATATGCGCGTCTCTAAAGGATAGATTTACATA 3060
Db 3001 ATTCAATCCAGATAGCAGTACGATTAACATATGCGCGTCTCTAAAGGATAGATTTACATA 3060
Qy 3061 TCGGAGACAAACCCGAGCAACATTTGTTCTGCAATGGAACAGCTCAGACCCGAGAC 3120
Db 3061 TCGGAGACAAACCCGAGCAACATTTGTTCTGCAATGGAACAGCTCAGACCCGAGAC 3120
Qy 3121 ACCGGGTCTATTTTGTGAAGCTTTACTTTCGGCTTCCCTGGTGTGCTTATTTGGGGC 3180
Db 3121 ACCGGGTCTATTTTGTGAAGCTTTACTTTCGGCTTCCCTGGTGTGCTTATTTGGGGC 3180
Qy 3181 CAAGGACCCCGGTACCGTCTCCAGTCTTAAGCCACACGACCCGAGCCGCGGACCA 3240
Db 3181 CAAGGACCCCGGTACCGTCTCCAGTCTTAAGCCACACGACCCGAGCCGCGGACCA 3240
Qy 3241 CCAACACCGGCGCCACCATCGCTGCGAGCCCTGCTCCCTGCGCCACAGAGCGGCTCGG 3300
Db 3241 CCAACACCGGCGCCACCATCGCTGCGAGCCCTGCTCCCTGCGCCACAGAGCGGCTCGG 3300
Qy 3301 CCAGCGGGGGGGCGAGTGACACAGAGGGGTGGAATTCGCGCTCGATCCCAAACTC 3360
Db 3301 CCAGCGGGGGGGCGAGTGACACAGAGGGGTGGAATTCGCGCTCGATCCCAAACTC 3360

Db 3301 CCAGCGGGGGGGCGAGTGACACAGAGGGGGCTGGACTTCGCGCTCGATCCCAAACTC 3360
Qy 3361 TGCTACCTCTGGATGGAATCTCTTCACTATATGTGTCACTTCTCACTGCTGTCTCTG 3420
Db 3361 TGCTACCTCTGGATGGAATCTCTTCACTATATGTGTCACTTCTCACTGCTGTCTCTG 3420
Qy 3421 AGAGTGAAGTTTACAGAGGCGCAGAGCCCCCGGTTACAGAGGGGCGAGAACGAGTCC 3480
Db 3421 AGAGTGAAGTTTACAGAGGCGCAGAGCCCCCGGTTACAGAGGGGCGAGAACGAGTCC 3480
Qy 3481 TATAACGAGCTCAATCTTAGGACGAAGAGAGGATACGATGTTTTGGCAAGAGAGCTGGC 3540
Db 3481 TATAACGAGCTCAATCTTAGGACGAAGAGAGGATACGATGTTTTGGCAAGAGAGCTGGC 3540
Qy 3541 CGGACCTCTGAGATGCGGGGAAAGCCGAGAGGAAACCTCTCAGAGAGGCTGTACAAAT 3600
Db 3541 CGGACCTCTGAGATGCGGGGAAAGCCGAGAGGAAACCTCTCAGAGAGGCTGTACAAAT 3600
Qy 3601 GAACTGCAAGAAAGATTAAGATGGCGAGGCTTACAGTGAAGATTGGATGAAAGGCGAGCGC 3660
Db 3601 GAACTGCAAGAAAGATTAAGATGGCGAGGCTTACAGTGAAGATTGGATGAAAGGCGAGCGC 3660
Qy 3661 CGGAGGGGCAAGGGGCAAGTGGCTTTTACAGGGTCTCAGTACAGCCACCAAGAGACCC 3720
Db 3661 CGGAGGGGCAAGGGGCAAGTGGCTTTTACAGGGTCTCAGTACAGCCACCAAGAGACCC 3720
Qy 3721 TACGACGCTTCACTATGTCAGGCTCTGCGCTTAACTCGACGCGCGCGGATCCG 3780
Db 3721 TACGACGCTTCACTATGTCAGGCTCTGCGCTTAACTCGACGCGCGCGGATCCG 3780
Qy 3781 GATTAGTCCAATTTGTTAAAGACAGGATATCAGTGGTTCAGGCTCTAGTTTGTACTCAAC 3840
Db 3781 GATTAGTCCAATTTGTTAAAGACAGGATATCAGTGGTTCAGGCTCTAGTTTGTACTCAAC 3840
Qy 3841 AATATCAACCAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900
Db 3841 AATATCAACCAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTA 3900
Qy 3901 GTCTCCAGAAAAGGGGGGAATGAAAGACCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960
Db 3901 GTCTCCAGAAAAGGGGGGAATGAAAGACCCACTGTAGGTTTGGCAAGCTAGCTTAAG 3960
Qy 3961 TAACGCCATTTTCAAGGCAATGAAAATACATACTGAGATAGAGAGTTTCAAGTCAAG 4020
Db 3961 TAACGCCATTTTCAAGGCAATGAAAATACATACTGAGATAGAGAGTTTCAAGTCAAG 4020
Qy 4021 GTTAGGAAACAGAGAGACAGAGGATATGCGGCAACAGGATATCTGTGGTAAGCGATTCC 4080
Db 4021 GTTAGGAAACAGAGAGACAGAGGATATGCGGCAACAGGATATCTGTGGTAAGCGATTCC 4080
Qy 4081 TGCCCGCTCAGGCGCAAGAACAGTTGGAAACAGGAGATATGGGCAACAGGATATCTG 4140
Db 4081 TGCCCGCTCAGGCGCAAGAACAGTTGGAAACAGGAGATATGGGCAACAGGATATCTG 4140
Qy 4141 TGGTAAGCAGTTCTGCTGCGCGCTCAGGCGCAAGAACAGATGGTCCCAGATCGGCTCC 4200
Db 4141 TGGTAAGCAGTTCTGCTGCGCGCTCAGGCGCAAGAACAGATGGTCCCAGATCGGCTCC 4200
Qy 4201 GCGCTCAGAGTCTTATAGAGAACCATCAGATGTTTTCAAGGGTGGCCCAAGGACCTGAAAT 4260
Db 4201 GCGCTCAGAGTCTTATAGAGAACCATCAGATGTTTTCAAGGGTGGCCCAAGGACCTGAAAT 4260
Qy 4261 GACCTGTGCTTATTTGAACTAACTCAATCAGTTCTGCTTCTGCTTCTGCTGCGCGCTT 4320
Db 4261 GACCTGTGCTTATTTGAACTAACTCAATCAGTTCTGCTTCTGCTTCTGCTGCGCGCTT 4320
Qy 4321 CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGGCGGCTCTCCGA 4380
Db 4321 CTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGGCGGCTCTCCGA 4380
Qy 4381 TAGACTCGTCCCGGCTACCGGTGTTCTCAATAAACCCTCTTTCAGATTGCATCCGACT 4440
Db 4381 TAGACTCGTCCCGGCTACCGGTGTTCTCAATAAACCCTCTTTCAGATTGCATCCGACT 4440

QY	4441	CGTGGTCTCGCTGTTCTCTGGGAGGGTCTCTCTGAGTGATTTGACTACCCGTCAGCGGGT	4500	QY	5521	GATACCAAGCGGTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTCTCCGACCCCTGCGCG	5580
DB	4441	CGTGGTCTCGCTGTTCTCTGGGAGGGTCTCTCTGAGTGATTTGACTACCCGTCAGCGGGT	4500	DB	5521	GATACCAAGCGGTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTCTCCGACCCCTGCGCG	5580
QY	4501	CTTTTCAGTTTCTCCCACTACACAGAGTCTCACTAAACATTCCTGATGTGCGCAGGACATC	4560	QY	5581	TTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC	5640
DB	4501	CTTTTCAGTTTCTCCCACTACACAGAGTCTCACTAAACATTCCTGATGTGCGCAGGACATC	4560	DB	5581	TTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC	5640
QY	4561	CGTCAGCCCGGTTTTGTTTATAATAAATGAAGACAGTGTTCCTTCAAGCCAGACT	4620	QY	5641	GCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTCCTCAAGCTGGCTGTGTGACAGAAC	5700
DB	4561	CGTCAGCCCGGTTTTGTTTATAATAAATGAAGACAGTGTTCCTTCAAGCCAGACT	4620	DB	5641	GCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGTCCTCAAGCTGGCTGTGTGACAGAAC	5700
QY	4621	ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCGACA	4680	QY	5701	CCCCCGTTTCAGCCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAGTCAACCCGG	5760
DB	4621	ACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCGACA	4680	DB	5701	CCCCCGTTTCAGCCCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAGTCAACCCGG	5760
QY	4681	CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCGTCTA	4740	QY	5761	TAAGACAGACTTATTCGCACTTGGCAGCAGCACTGGTAAACAGGATTTAGCAGAGGAGT	5820
DB	4681	CGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATCTATCTCGTCTA	4740	DB	5761	TAAGACAGACTTATTCGCACTTGGCAGCAGCACTGGTAAACAGGATTTAGCAGAGGAGT	5820
QY	4741	TTTTGCTGCTTCTCGATTTTAAATTTCTAGTTTGCACCTCCCTTCTTGAGACACGCG	4800	QY	5821	ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAAGA	5880
DB	4741	TTTTGCTGCTTCTCGATTTTAAATTTCTAGTTTGCACCTCCCTTCTTGAGACACGCG	4800	DB	5821	ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACCGCTACACTAGAAGA	5880
QY	4801	GATTGACAGAGTAAATACTCTGAGGCGAGGCTCTGTGAAAAGGTTGCTGGGCTCAG	4860	QY	5881	CAGTATTTGGTATCTGCTGAAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTAGCT	5940
DB	4801	GATTGACAGAGTAAATACTCTGAGGCGAGGCTCTGTGAAAAGGTTGCTGGGCTCAG	4860	DB	5881	CAGTATTTGGTATCTGCTGAAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTAGCT	5940
QY	4861	TGTGAGATTTGCCATAAAGGGGTCCTGCCCTCTGTGTACAGACAGATCGGAATCTAGA	4920	QY	5941	CTTGATTCGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTGTTTGCAGAGCAGAGA	6000
DB	4861	TGTGAGATTTGCCATAAAGGGGTCCTGCCCTCTGTGTACAGACAGATCGGAATCTAGA	4920	DB	5941	CTTGATTCGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTGTTTGCAGAGCAGAGA	6000
QY	4921	GTGCATCTACAGTCCCGGTTCCGGGGTCTGATCTCAGGGATCTTTGCGCTAGAG	4980	QY	6001	TTACCGCGAGAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACGGGGTCTGAGC	6060
DB	4921	GTGCATCTACAGTCCCGGTTCCGGGGTCTGATCTCAGGGATCTTTGCGCTAGAG	4980	DB	6001	TTACCGCGAGAAAAAGGATCTCAAGAGATCCCTTTGATCTTTCTACGGGGTCTGAGC	6060
QY	4981	ATCCTCTACCGCGAGCGATCGTGGCGGTACCGAGCTCGAATCTGTAATCATGTCTAT	5040	QY	6061	CTCAGTGAAACGAAACTCACTTAAAGGATTTTGGTCTATGAGATTAATAAAGGATCT	6120
DB	4981	ATCCTCTACCGCGAGCGATCGTGGCGGTACCGAGCTCGAATCTGTAATCATGTCTAT	5040	DB	6061	CTCAGTGAAACGAAACTCACTTAAAGGATTTTGGTCTATGAGATTAATAAAGGATCT	6120
QY	5041	AGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCACACATACAGAGCGGAA	5100	QY	6121	TCACCTAGATCTTTTAAATTAAGATTTTAAATTAAGATTTTAAATTAAGATTTATATAGT	6180
DB	5041	AGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATCCACACATACAGAGCGGAA	5100	DB	6121	TCACCTAGATCTTTTAAATTAAGATTTTAAATTAAGATTTTAAATTAAGATTTATATAGT	6180
QY	5101	GCATAAAGTGAAGCTGGGTGCTTAATGATGAGTAACTCACTCAATTAATTTGGTTGC	5160	QY	6181	AACTTTGCTGACAGTTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTC	6240
DB	5101	GCATAAAGTGAAGCTGGGTGCTTAATGATGAGTAACTCACTCAATTAATTTGGTTGC	5160	DB	6181	AACTTTGCTGACAGTTTACCAATGCTTAATCAGTGAGGACCTATCTCAGCGATCTGTC	6240
QY	5161	GCTCACTGCCCTTTCCAGTCCGGAAACCTGTGCGCAGCTGCATTAATGAATCGGCG	5220	QY	6241	TATTTTCGTTTCACTAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG	6300
DB	5161	GCTCACTGCCCTTTCCAGTCCGGAAACCTGTGCGCAGCTGCATTAATGAATCGGCG	5220	DB	6241	TATTTTCGTTTCACTAGTTGCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG	6300
QY	5221	AACCGCGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTTCCGTTTCTCGCTCACTGACT	5280	QY	6301	GCTTACCATCTGGCCCCAGTGTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG	6360
DB	5221	AACCGCGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTTCCGTTTCTCGCTCACTGACT	5280	DB	6301	GCTTACCATCTGGCCCCAGTGTCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG	6360
QY	5281	CGTGGCTCGGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAAAGCGGTAATAC	5340	QY	6361	ATTTATCAGCAATAAACACGAGCCGGAAGGCGGAGGAGTGGTCTCTGCAACTT	6420
DB	5281	CGTGGCTCGGTTCGGCTGCGGCGAGCGGTATCAGTCACTCAAAAGCGGTAATAC	5340	DB	6361	ATTTATCAGCAATAAACACGAGCCGGAAGGCGGAGGAGTGGTCTCTGCAACTT	6420
QY	5341	GTTTATCCACAGATCAGGGGATACCGAGGAAAGCAATGTGACAAAAGGCCAGCAA	5400	QY	6421	TATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCGAG	6480
DB	5341	GTTTATCCACAGATCAGGGGATACCGAGGAAAGCAATGTGACAAAAGGCCAGCAA	5400	DB	6421	TATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAGAGTAAGTAGTTCCGCGAG	6480
QY	5401	AGGCCAGGAACCGTAAAAAGGCGGTTCTGCGGTTTTTCCATAGGCTTCGCCCCCTTG	5460	QY	6481	TTAATAGTTTGGCAACCGTTGTCATTTGCTACAGGCTCGTGGTGTCACTGCTCGT	6540
DB	5401	AGGCCAGGAACCGTAAAAAGGCGGTTCTGCGGTTTTTCCATAGGCTTCGCCCCCTTG	5460	DB	6481	TTAATAGTTTGGCAACCGTTGTCATTTGCTACAGGCTCGTGGTGTCACTGCTCGT	6540
QY	5461	ACGAGCATCAAAAAATCAGCGCTCAAGTCAAGAGTGGGAAACCCGACAGGACTATAAA	5520	QY	6541	TGGTATGGCTTCAATTCAGTCCGTTCCCAAGATCAAGGGGAGTTACATGATCCCCAT	6600
DB	5461	ACGAGCATCAAAAAATCAGCGCTCAAGTCAAGAGTGGGAAACCCGACAGGACTATAAA	5520	DB	6541	TGGTATGGCTTCAATTCAGTCCGTTCCCAAGATCAAGGGGAGTTACATGATCCCCAT	6600
				QY	6601	GTTGTGCAAAAAAGCGGTTAGCTCCTCGGTCCTCCGATCGTTGTGTCAGAAAGTAAGTTGCG	6660

6601 GTTGTGCAAAAAGCGGTAGTCTCTCGGTCTCGATCGTTGTGCAAGTAGTTGGC 6660
6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGATTAATTTCTTACTGTCAAGCCATC 6720
6661 CGCAGTGTATCACTCATGTTATGGCAGCACTGATTAATTTCTTACTGTCAAGCCATC 6720
6721 CGTAGAGTCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCATTTCTGAGATAGTGAT 6780
6721 CGTAGAGTCTTTTCTGTGACTGGTGAAGTACTCAACCAAGTCATTTCTGAGATAGTGAT 6780
6781 GCGGCGACCGAGTTGCTCTTGTCCCGCGCTCAATACGGGATAATACCGGCCACATAGCAG 6840
6781 GCGGCGACCGAGTTGCTCTTGTCCCGCGCTCAATACGGGATAATACCGGCCACATAGCAG 6840
6841 AACTTTAAAGTCTCATCATTTGAAAAAGTTTTCGGGGGAAAACTCTCAAGGATCTT 6900
6841 AACTTTAAAGTCTCATCATTTGAAAAAGTTTTCGGGGGAAAACTCTCAAGGATCTT 6900
6901 ACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTCAGCATC 6960
6901 ACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGCACCACTGATCTTCAGCATC 6960
6961 TTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCGCAAAAAA 7020
6961 TTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGAAAGGCAAAATGCGCAAAAAA 7020
7021 GGGATAAGGCGCACACGGAATTTGAATTAATCTCATACTCTTCTTTTCAATATTATTG 7080
7021 GGGATAAGGCGCACACGGAATTTGAATTAATCTCATACTCTTCTTTTCAATATTATTG 7080
7081 AAGCAATTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGAATTTAGAAAAA 7140
7081 AAGCAATTATCAGGGTTATTGTTCTCATGAGCGGATACATATTGAATTTAGAAAAA 7140
7141 TAAACAATAGGGTTCCGCGCACATTTCCCGAAAAGTGCCACTGACGTTCTAAGAAC 7200
7141 TAAACAATAGGGTTCCGCGCACATTTCCCGAAAAGTGCCACTGACGTTCTAAGAAC 7200
7201 CATTTATATCATGACATTAACCTATATAAATAGGCGTATCACAGGCGCTTCTGTCGCG 7260
7201 CATTTATATCATGACATTAACCTATATAAATAGGCGTATCACAGGCGCTTCTGTCGCG 7260
7261 CGCTTTCGGTGATGACGCGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGC 7320
7261 CGCTTTCGGTGATGACGCGTGAACCTCTGACACATGACGCTCCCGGAGACGGTCAACAGC 7320
7321 TTGTCGTAAAGGATGCGGAGCAGACAGCCGCTCAGGCGCGGTGAGCGGGTGTGG 7380
7321 TTGTCGTAAAGGATGCGGAGCAGACAGCCGCTCAGGCGCGGTGAGCGGGTGTGG 7380
7381 CGGGTGTCCGGCTGGCTTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCA 7440
7381 CGGGTGTCCGGCTGGCTTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCA 7440
7441 TATCGGTGTGAATAACCGCACAGATGCGTAAAGGAGAAAAATACCGCATFCAGCGGCCATTTC 7500
7441 TATCGGTGTGAATAACCGCACAGATGCGTAAAGGAGAAAAATACCGCATFCAGCGGCCATTTC 7500
7501 GCCATTACGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACG 7560
7501 GCCATTACGGCTGCGCAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACG 7560
7561 CCAGTCGCAAGAGGGGATGCTGCAAGCGGATTAAGTTGGGTAAAGCGGAGGGTTTC 7620
7561 CCAGTCGCAAGAGGGGATGCTGCAAGCGGATTAAGTTGGGTAAAGCGGAGGGTTTC 7620
7621 CCAGTCACGAGTTGTAAAAACGACGGCCAGTGCC 7654
7621 CCAGTCACGAGTTGTAAAAACGACGGCCAGTGCC 7654

US-10-006-771A-1
; Sequence 1, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2428)..(3759)
; OTHER INFORMATION: Chimeric IgTCR sequence contained in retroviral vector. Retrovi
; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv
; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent
; OTHER INFORMATION: ion. (pertinent to Figure 3.)
US-10-006-771A-1

Query Match 100.0%; Score 7654; DB 13; Length 7654;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCTTGATGCTGCGAGTGCAGTCTAGGCACATATAAGAAAAACATTAACCAAGCT 60
Db 1 AAGCTTGATGCTGCGAGTGCAGTCTAGGCACATATAAGAAAAACATTAACCAAGCT 60
Qy 61 GCACCCGAGCAGTGAAGAAGCCGTTAAACCGTTTGTAAATAAATGAAATTTT 120
Db 61 GCACCCGAGCAGTGAAGAAGCCGTTAAACCGTTTGTAAATAAATGAAATTTT 120
Qy 121 AGAGTCATTTCTTTGGTAGAAAGTACATTTGGCAGCTAAAGAGGCCAAAGCAATCTGTG 180
Db 121 AGAGTCATTTCTTTGGTAGAAAGTACATTTGGCAGCTAAAGAGGCCAAAGCAATCTGTG 180
Qy 181 GAAAGCCAGCTGGGAGCCAGCAGTTGTCATCCCTCTCTGGCGTGTACTAAGGGTTT 240
Db 181 GAAAGCCAGCTGGGAGCCAGCAGTTGTCATCCCTCTCTGGCGTGTACTAAGGGTTT 240
Qy 241 CTTAAATTTGTGGTTCTTAATCTTCCAGAGGGTTTGTCTCATTTCACTTCCACTTCGGTG 300
Db 241 CTTAAATTTGTGGTTCTTAATCTTCCAGAGGGTTTGTCTCATTTCACTTCCACTTCGGTG 300
Qy 301 CACAATCTTGGACGCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTTAGGC 360
Db 301 CACAATCTTGGACGCGGATTTACTGTCTTAGCATCTATCGGTGGCCCTTCGATTTAGGC 360
Qy 361 TGAACCTTGAGCCCACTTCTTACGTTGTTAAAGAGAGCACAAGCACCAGAAAGGCTGA 420
Db 361 TGAACCTTGAGCCCACTTCTTACGTTGTTAAAGAGAGCACAAGCACCAGAAAGGCTGA 420
Qy 421 CCGGCGAGACCTGTGGGCAATTTTAAACAAGGCCCTCTGGTCTGTGGAGGAGGCTTA 480
Db 421 CCGGCGAGACCTGTGGGCAATTTTAAACAAGGCCCTCTGGTCTGTGGAGGAGGCTTA 480
Qy 481 CATAAGGTGCAAAATTAGAAATATAAATAAAGCCCATATCAATTTGTCTATCTTTT 540
Db 481 CATAAGGTGCAAAATTAGAAATATAAATAAAGCCCATATCAATTTGTCTATCTTTT 540
Qy 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAGCCCAAT 600
Db 541 AGCTCAAGTTTGAAGACCCCACTGTAGGTTTGGCAAGCTAGCTTAAGTAAGCCCAAT 600
Qy 601 TTGCAAGGCATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCAAGGTTAGGAACA 660
Db 601 TTGCAAGGCATGGAATAATACATACTGAGAAATAGAGAGTTTCAGATCAAGGTTAGGAACA 660

[illegible]

QY 5041 AGCTGTTTCTGTGAAATTTGTTATCCGCTCACAATTCACCAACATACGAGCGGAA 5100
 Db |||||
 QY 5041 AGCTGTTTCTGTGAAATTTGTTATCCGCTCACAATTCACCAACATACGAGCGGAA 5100
 Db |||||
 QY 5101 GCATAAAGTGTAAAGCTCGGGTGCTTAATAGTGTAGCTAACTCAATTAATTTGGTTGC 5160
 Db |||||
 QY 5101 GCATAAAGTGTAAAGCTCGGGTGCTTAATAGTGTAGCTAACTCAATTAATTTGGTTGC 5160
 Db |||||
 QY 5161 GCTCACTGCCCTTTTCCAGTCCGGAAACCTGTGTCGAGCTGCATTAATGAATCGGCC 5220
 Db |||||
 QY 5161 GCTCACTGCCCTTTTCCAGTCCGGAAACCTGTGTCGAGCTGCATTAATGAATCGGCC 5220
 Db |||||
 QY 5221 AACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCTCGCTCACTGACT 5280
 Db |||||
 QY 5221 AACCGCGGGGAGAGCGGTTTGGGTATTGGCGCTCTTCCGCTTCTCGCTCACTGACT 5280
 Db |||||
 QY 5281 CGCTGCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATAC 5340
 Db |||||
 QY 5281 CGCTGCGCTCGGTTCGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATAC 5340
 Db |||||
 QY 5341 GGTATCCACAGAAATCAGGGGATACGCAAGGAAAGAAATGATGAGCAAAAGGCCAGCAA 5400
 Db |||||
 QY 5341 GGTATCCACAGAAATCAGGGGATACGCAAGGAAAGAAATGATGAGCAAAAGGCCAGCAA 5400
 Db |||||
 QY 5401 AGGCCAGGAACCGTAAAGGCGCGTTGCTGCGGTTTTTCCATAGGCTCGGCCCGCTG 5460
 Db |||||
 QY 5401 AGGCCAGGAACCGTAAAGGCGCGTTGCTGCGGTTTTTCCATAGGCTCGGCCCGCTG 5460
 Db |||||
 QY 5461 ACAGCATCACAAAATCAGCTCAAGTCAAGGTGGCGAAACCCGACAGGACTATAAA 5520
 Db |||||
 QY 5461 ACAGCATCACAAAATCAGCTCAAGTCAAGGTGGCGAAACCCGACAGGACTATAAA 5520
 Db |||||
 QY 5521 GATACAGCGGTTTCCCGCTGAAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTCGCGC 5580
 Db |||||
 QY 5521 GATACAGCGGTTTCCCGCTGAAAGCTCCCTCGTGGGCTCTCTGTTCCGACCTCGCGC 5580
 Db |||||
 QY 5581 TTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCGTGGCTTCTCATAGCTCAC 5640
 Db |||||
 QY 5581 TTACCGGATACCTGTCGCGCTTTTCCCTTCGGGAAGCGTGGCTTCTCATAGCTCAC 5640
 Db |||||
 QY 5641 GCTGTAGGTATCTCAGTTCGGTGTAGTGTGTTCCGCTTCCGCTGCGCTGTTGCAAGT 5700
 Db |||||
 QY 5641 GCTGTAGGTATCTCAGTTCGGTGTAGTGTGTTCCGCTTCCGCTGCGCTGTTGCAAGT 5700
 Db |||||
 QY 5701 CCCCCTTCAGCCGACCGCTGAGCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGG 5760
 Db |||||
 QY 5701 CCCCCTTCAGCCGACCGCTGAGCTTATCCGGTAACTATCGTCTTGTAGTCAACCCGG 5760
 Db |||||
 QY 5761 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGT 5820
 Db |||||
 QY 5761 TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGT 5820
 Db |||||
 QY 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAAGA 5880
 Db |||||
 QY 5821 ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAAGA 5880
 Db |||||
 QY 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATAAGAGTTGGTAGCT 5940
 Db |||||
 QY 5881 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATAAGAGTTGGTAGCT 5940
 Db |||||
 QY 5941 CTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTGTTGCAAGCAGCAGA 6000
 Db |||||
 QY 5941 CTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTGTTGCAAGCAGCAGA 6000
 Db |||||
 QY 6001 TTACGGCGCAAAAAAGGATCTCAAGAAAGATCTTTTGAATCTTTTCTACGGGTCTGACG 6060
 Db |||||
 QY 6001 TTACGGCGCAAAAAAGGATCTCAAGAAAGATCTTTTGAATCTTTTCTACGGGTCTGACG 6060
 Db |||||
 QY 6061 CTCAGTGGAAACGAAACTCAGGTTAAGGATTTTGGTCAAGAGATTAACAAAAGATCT 6120
 Db |||||
 QY 6061 CTCAGTGGAAACGAAACTCAGGTTAAGGATTTTGGTCAAGAGATTAACAAAAGATCT 6120
 Db |||||

QY 6121 TCACCTAGATCCTTTTAAATTTAAATGAAGTTTTTAAATCAATCTAAAGTATATATAGT 6180
 Db |||||
 QY 6121 TCACCTAGATCCTTTTAAATTTAAATGAAGTTTTTAAATCAATCTAAAGTATATATAGT 6180
 Db |||||
 QY 6181 AAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTC 6240
 Db |||||
 QY 6181 AAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCACTATCTCAGCGATCTGTC 6240
 Db |||||
 QY 6241 TATTTTCGTTTCACTCCAGTTCCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300
 Db |||||
 QY 6241 TATTTTCGTTTCACTCCAGTTCCTGACTCCCGCTCGTGTAGATACTACGATACGGGAGG 6300
 Db |||||
 QY 6301 GCTTACCATCTGGGCCCCAGTGTGCAATGATATACCCGAGACCCACGCTCACCGGCTCCAG 6360
 Db |||||
 QY 6301 GCTTACCATCTGGGCCCCAGTGTGCAATGATATACCCGAGACCCACGCTCACCGGCTCCAG 6360
 Db |||||
 QY 6361 ATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCGAACTT 6420
 Db |||||
 QY 6361 ATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTCGAACTT 6420
 Db |||||
 QY 6421 TATCCGCTCCCATCCAGTCTATTAAATGTTGCGGGGAAGCTAGAGTAAGTAGTTCGCCAG 6480
 Db |||||
 QY 6421 TATCCGCTCCCATCCAGTCTATTAAATGTTGCGGGGAAGCTAGAGTAAGTAGTTCGCCAG 6480
 Db |||||
 QY 6481 TTAATAGTTTGGCAACCGTTGTCATACAGGCTCGTGGTGTCAACGCTCGCTCGTT 6540
 Db |||||
 QY 6481 TTAATAGTTTGGCAACCGTTGTCATACAGGCTCGTGGTGTCAACGCTCGCTCGTT 6540
 Db |||||
 QY 6541 TGGTATGCTTCAATTCAGTCCGCTTCCAAACGATCAAGGGGAGTTAATGATTCGCCAT 6600
 Db |||||
 QY 6541 TGGTATGCTTCAATTCAGTCCGCTTCCAAACGATCAAGGGGAGTTAATGATTCGCCAT 6600
 Db |||||
 QY 6601 GTTGTGCAAAAAGCGGTTAGCTCCTCGTCCGATCGTTGTCAGAAAGTAAAGTGGC 6660
 Db |||||
 QY 6601 GTTGTGCAAAAAGCGGTTAGCTCCTCGTCCGATCGTTGTCAGAAAGTAAAGTGGC 6660
 Db |||||
 QY 6661 GCGAGTGTATCAGTTCATGAGTATGCGAGCACTGATTAATCTCTTACTGTCAAGCATC 6720
 Db |||||
 QY 6661 GCGAGTGTATCAGTTCATGAGTATGCGAGCACTGATTAATCTCTTACTGTCAAGCATC 6720
 Db |||||
 QY 6721 GGTAAAGTGTCTTGTGACTGTGGTGTGATCTCAACCAAGTCATTTCTGAGAAATAGTAT 6780
 Db |||||
 QY 6721 GGTAAAGTGTCTTGTGACTGTGGTGTGATCTCAACCAAGTCATTTCTGAGAAATAGTAT 6780
 Db |||||
 QY 6781 GCGCGACCGAGTTCCTCTTGGCCCGGCTCAATAGCGGATTAATCCGCGCCACATAGCAG 6840
 Db |||||
 QY 6781 GCGCGACCGAGTTCCTCTTGGCCCGGCTCAATAGCGGATTAATCCGCGCCACATAGCAG 6840
 Db |||||
 QY 6841 AACTTTAAAGTGTCTCATCTTGGAAAAACGTTTCTTGGGGGCGAAAACTCTCAAGGATCTT 6900
 Db |||||
 QY 6841 AACTTTAAAGTGTCTCATCTTGGAAAAACGTTTCTTGGGGGCGAAAACTCTCAAGGATCTT 6900
 Db |||||
 QY 6901 ACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATC 6960
 Db |||||
 QY 6901 ACCGCTGTGAGATCCAGTTCGATGTAAACCACTCGTGACCCCACTGATCTTTCAGCATC 6960
 Db |||||
 QY 6961 TTTTACTTTTCCAGCGGTTCTGGGTGAGCAAAAACAGAGGCAAAAATCCGCGCAAAAA 7020
 Db |||||
 QY 6961 TTTTACTTTTCCAGCGGTTCTGGGTGAGCAAAAACAGAGGCAAAAATCCGCGCAAAAA 7020
 Db |||||
 QY 7021 GGGATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATATTG 7080
 Db |||||
 QY 7021 GGGATAAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATATTG 7080
 Db |||||
 QY 7081 AAGCAATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA 7140
 Db |||||
 QY 7081 AAGCAATTTATCAGGGTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAA 7140
 Db |||||
 QY 7141 TAAACAAATAGGGTTCCCGCACATTTCCCGGAAAGTGGCCACTGAGCTCTAGAAAC 7200
 Db |||||
 QY 7141 TAAACAAATAGGGTTCCCGCACATTTCCCGGAAAGTGGCCACTGAGCTCTAGAAAC 7200
 Db |||||
 QY 7201 CATTATTATCATGACATTAACCTATAAAAAATAGGGGTATACAGAGGCCCTTTCTGCTCGC 7260
 Db |||||

Db 7201 CATTATATCATGACATTAACCTATATAAATAGGGCTATCAGAGGCCCTTCTGCTCGC 7260
Qy GCGTTTCGGTGATGACGCTGAAACCTCTGACACATGACGATCCCGGAGACGGTCAACAGC 7320
Db GCGTTTCGGTGATGACGCTGAAACCTCTGACACATGACGATCCCGGAGACGGTCAACAGC 7320
Qy TTGCTCTGAAGCGGATGCCGGAGCAGACAGCCGCTCAGGCGCGCTCAGCGGGGTGTTGG 7380
Db TTGCTCTGAAGCGGATGCCGGAGCAGACAGCCGCTCAGGCGCGCTCAGCGGGGTGTTGG 7380
Qy CCGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGATGACCA 7440
Db CCGGTGTCCGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGACTGAGATGACCA 7440
Qy TATCGGTGTGAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCCATTC 7500
Db TATCGGTGTGAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCCATTC 7500
Qy GCCATTGAGGCTGGCAACTGTTGGGAAGGGGATCGGTGCGGGGCTCTTCGCTATTACG 7560
Db GCCATTGAGGCTGGCAACTGTTGGGAAGGGGATCGGTGCGGGGCTCTTCGCTATTACG 7560
Qy CCAGCTGGCGAAAGGGGATGCTGCAAGCGATTAAAGTTGGGTAAACGCCAGGTTTC 7620
Db CCAGCTGGCGAAAGGGGATGCTGCAAGCGATTAAAGTTGGGTAAACGCCAGGTTTC 7620
Qy CCAGTACAGCGTTGTATAACGACGCCCGCAGTGCC 7654
Db CCAGTACAGCGTTGTATAACGACGCCCGCAGTGCC 7654

RESULT 3

US-10-913-288-1
; Sequence 1, Application US/10913288
; Publication No. US20050075492A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, CHANG-ZHENG
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR EXPRESSION OF microRNA's
; FILE REFERENCE: W00571.70009 US
; CURRENT APPLICATION NUMBER: US/10/913,288
; CURRENT FILING DATE: 2004-08-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 6963
; TYPE: DNA
; ORGANISM: Retroviral
US-10-913-288-1

Query Match 51.0%; Score 3900; DB 19; Length 6963;
Best Local Similarity 75.8%; Pred. No. 0;
Matches 5394; Conservative 0; Mismatches 1130; Indels 596; Gaps 22;
Qy 552 TGAAGACCCCACTGTAGTTGGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGCAT 611
Db 1 TGAAGACCCCACTGTAGTTGGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGCAT 60
Qy 612 GGAATAACATACTGAGAAATAGAGAGTTTCAATCAAGTTAGGAACAGAGACAGCA 671
Db 61 GGAATAACATACTGAGAAATAGAGAGTTTCAATCAAGTTAGGAACAGAGACAGCA 119
Qy 672 GAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTCCCGCTCAGGCGCAAGAA 731
Db 120 ----- 119
Qy 732 CAGTTGGAAACAGGAAATATCGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTCCCGC 791
Db 120 -----AGAAATGGGCCAAACAGGATATCTGTGTAAGCAGTTCTCTGCCCCG 166
Qy 792 GCTCAGGCCCAAGAACAGATGGTCCCAGATGCGGTCCCGCCCTCAGCAGTTTCTAGAGA 851

Db 167 GCTCAGGCCCAAGAACAGATGGTCCCAGATGCGGTCCCGCCCTCAGCAGTTTCTAGAGA 226
Qy ACCATCAGATGTTTCCAGGGTGCCCAAGAGACTTG- AAATGACCTGTGCTTATTGAA 910
Db ACCATCAGATGTTTCCAGGGTGCCCAAGAGACTTGAATAATGACCTGTGCTTATTGAA 286
Qy CTAAACCAATCAGTTTCGCTTCTGCTTCTGTTTCGGCGGCTTCTGCTCCCGAGCTCAATAA 970
Db CTAAACCAATCAGTTTCGCTTCTGCTTCTGTTTCGGCGGCTTCTGCTCCCGAGCTCAATAA 346
Qy AAGAGCCCAACAAACCCCTCACTCGCGCGGCCAGTCTCTCGATAGACTGCGTCCCGGGTA 1030
Db AAGAGCCCAACAAACCCCTCACTCGCGCGGCCAGTCTCTCGATAGACTGCGTCCCGGGTA 406
Qy CCGTATTTCCCAATAAAGCCTCTTGTGTTTGGCATTCGGAATCGTGGATCTGCTGATCTT 1090
Db CCGTATTTCCCAATAAAGCCTCTTGTGTTTGGCATTCGGAATCGTGGATCTGCTGATCTT 466
Qy GGGAGGGTCTCTCAGATTGATTGACTGCCACCTCGGGGGTCTTTCAATTTGGAGGTTCC 1150
Db GGGAGGGTCTCTCAGATTGATTGACTGCCACCTCGGGGGTCTTTCAATTTGGAGGTTCC 526
Qy ACCGAGATTTGGAGACCCCTGCCCCAGGACCAACCGACCCCGCGGGAGGTAAAGCTGG 1210
Db ACCGAGATTTGGAGACCCCTGCCCCAGGACCAACCGACCCCGCGGGAGGTAAAGCTGG 586
Qy CCAGCAACTTATCTGCTCTGTC-----CGATTGTCTAGTCTCTATGACTGATTTTA 1262
Db CCAGCGTCTGTTCTGCTCTGCTCTGCTCTGTTGTCGGTCTTGTGCGGCATCTAATGTT 646
Qy TGGCCTCGCTCGGTACTAGTACTAGTCTGTAAGTCTGTTGCGGACCCGCTGCTGTA 1322
Db TGGCCTCGCTCGGTACTAGTACTAGTCTGTAAGTCTGTTGCGGACCCGCTGCTGTA 706
Qy CTGACGAGTTTCGAAACACCCCGCGCAACCTCGGAGACGCTCCAGGAGCTTCGCGGGCC 1382
Db CTGACGAGTTTCGAAACACCCCGCGCAACCTCGGAGACGCTCCAGGAGCTTCGCGGGCC 766
Qy GTTTTGTGCGCCGACCTGAGTCTCTAAATCCCGATCGTTTGAAGACTCTTTGGTGACCC 1442
Db GTTTTGTGCGCCGACCTGAGGAGGAGTCTGATGTGGAATCCGACCC----- 814
Qy CCCTTAGAGAGGGATATGTTGTTCTGTTAGGAGACGAGAACCTTAAACAGTTCCCGCT 1502
Db -----CGTCAGGATATGTTTCTGTTAGGAGACGAGAACCTTAAACAGTTCCCGCT 867
Qy CCGTCTGAAATTTTGTCTTGGGACCGAAGCGCGCGCGCTTGTGT---CTGC 1559
Db CCGTCTGAAATTTTGTCTTGGGACCGAAGCGCGCGCGCTTGTGTCTGCTGCGAGCGC 927
Qy TGCAAGATCGTTCTGTTGTTCTGCTGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTG 1619
Db TGCAAGATCGTTCTGTTGTTCTGCTGCTGATCTGTTGTTGTTGTTGTTGTTGTTGTTG 987
Qy GGGCGGGCTAGACTGTTACATCCCTTAAGTTTGAACCTTAGTCTAGTCTAGTCACTGAAAGATGTC 1679
Db GGGC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGTCTAGTCACTGAAAGATGTC 1041
Qy GAGCGGATCGCTCAACACAGTCTGTAAGTCTCAAGAGAGACGCTTGGTTTACCTTCTGTC 1739
Db GAGCGGATCGCTCAACACAGTCTGTAAGTCTCAAGAGAGACGCTTGGTTTACCTTCTGTC 1101
Qy TCTCAGAAATGGCCAAACCTTTAACTCGGATGCGCGGAGACGCGACCTTTAAACCGAGAC 1799
Db TCTCAGAAATGGCCAAACCTTTAACTCGGATGCGCGGAGACGCGACCTTTAAACCGAGAC 1161
Qy CTATATCCAGGTTAAGATCAAGTCTTTTCACTGCGCGGATGAGACATCCAGACAG 1859
Db CTATATCCAGGTTAAGATCAAGTCTTTTCACTGCGCGGATGAGACATCCAGACAG 1221
Qy GTCCCTTACATCTGAGCTCGGAGCTTGGCTTTTGAACCCCTTCTGCTGCTGCTGCTGCTG 1919
Db GTCCCTTACATCTGAGCTCGGAGCTTGGCTTTTGAACCCCTTCTGCTGCTGCTGCTGCTG 1281

Qy	1920	TTTGTGACACCTAAGCGCTCGCGCTCTCTTCTCTCCATCCGCCCGCTCTCTCCCGCTTGAA	1979
Db	1282	TTTGTACACCTTAAGCGCTCGCGCTCTCTTCTCTCATCCGCCCGCTCTCTCCCGCTTGAA	1341
Qy	1980	CCTCCTCGTTGACGCCCGCGCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGC	2039
Db	1342	CCTCCTCGTTGCGACCCCGCGCTCGATCCTCCCTTTATCCAGCCCTCACTCCTTCTCTAGGC	1401
Qy	2040	GCCCCCATATGCCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCT	2099
Db	1402	GC-----CGGAATTAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCT	1451
Qy	2100	GACCTGACATGACAAGAGTTACTAAACAGCCCTCTCTTCCAAGTCTACTTTACAGGCTTC	2158
Db	1452	GACCTGACATGACAAGAGTTACTTAAACAGCCCTCTCTTCCAAGTCTACTTTACAGGCTTC	1511
Qy	2159	TACTTTAGTCAGACAGAAAGTCGAGACCTCTGCGCGCAGCCTTACCAGAACAACTCGAC	2218
Db	1512	TACTTTAGTCAGACAGAAAGTCGAGACCTCTGCGCGCAGCCTTACCAGAACAACTCGAC	1571
Qy	2219	CGACCGGTGGTCACTCAACCTTACCAGTCCGGGACACAGTGTGGGTCCGCGCACACAG	2278
Db	1572	CGACCGGTGGTCACTCAACCTTACCAGTCCGGGACACAGTGTGGGTCCGCGCACACAG	1631
Qy	2279	ACTTAAGAACCTTAGAACCTCGCTGGAAGAGACCTTTACACAGTCTGCTGACGACCCCGACC	2338
Db	1632	ACTTAAGAACCTTAGAACCTCGCTGGAAGAGACCTTTACACAGTCTGCTGACGACCCCGACC	1691
Qy	2339	GCCCTCAAAGTAGACGCGCATCGCAGCTTGGATACACGCCGCCACAGTGAAAGGTCGCCGAC	2398
Db	1692	GCCCTCAAAGTAGACGCGCATCGCAGCTTGGATACACGCCGCCACAGTGAAAGGTCGCCGAC	1751
Qy	2399	CCCGGGGTGGACCATCCTCTPAGACTGCCATGGGATGGAGCTGATCATCTCTTTTGG	2458
Db	1752	CCCGGGGTGGACCATCCTCTPAGACTGCC--GGATCAATTCTACCGGGTAGGGGAGCGC	1810
Qy	2459	TAGCAACAGCTACAGGTGTCCACTCCGACATCCAGTGAACCCAGAGCCCAAGCAGCTGA	2518
Db	1811	TTTTTCCAAAGCAGTCTGGAGCATGCGCTTTAGCGGCCCGCGCTGGGCATCTTGGCGCTACA	1870
Qy	2519	GCGCCAGCGTGGGTGACAGATGACCATCACCTCTTAAGGCGCAGTCAAGATGTGGGTACTT	2578
Db	1871	CAAGTGGCTCTGSCCTCGCACATTTCCACATCCACCGGTAGGGGCCAACCGGCTCCGT	1930
Qy	2579	CTGTAGCTTTGGTACCAAGAGAGCCAGGTAAAGGCTCCAAAGCTGCTGATCTACTGTGACAT	2638
Db	1931	TCTTTGGTGGCCCCCTTCGCGCCACTTCTACTCTCTCCCTTAGTCAGGAAGTCTCCCCCGC	1990
Qy	2639	CCACCCGCGACACTGGGTGCCAAGCAGATTCACGCGGTAGCGGTAGCGGTACCGACTTCA	2698
Db	1991	CCCCGCGAGCTCGCTCTGTGAGGACGTGACAAATGGAAGTAGACACGCTCTCACTAGTCTCG	2050
Qy	2699	CCTTTCACCATCAGCAGCGCTCCAGCCAGGAGACATTCGCCACTACTACTGTGCCA--GCAATAT	2757
Db	2051	TGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCTTTGGGGCAGCGGCCAAT	2110
Qy	2758	AGCCTCTATCGGTCTGTCGGCCAAAGGACCAAGGTGGAATTCAAAACGAGGTGGCTCAGGA	2817
Db	2111	AGCAGCTTTGGCTCCTTTCGCTTTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGGGC	2170
Qy	2818	TCGGGTGGATCCGGCTCTGTTGGCTCAGGATCCGAGGTCCAACTGGTGGAGAGCGGTGGA	2877
Db	2171	GGGCTCAGGGCGGGCTCAGGGGC-----GGGGCGGGCGCCCGAAGGTCTCTCGGAGG	2223
Qy	2878	GGTGTGTGCAACTGTGCCGCTCCCTCGCCTGTCTGTCTCCGATCTGGCTTCGATTTTC	2937
Db	2224	CCCGGCATCTGACGCTTCAAAAGCGCATCTGTGCGCGCTGTTCTCTCTCTTCTCTCATC	2283
Qy	2938	ACCACATATTGGATGAGTTGGTGGAGACAGGCACCTGGAAGAGTCTTTAGTGGATTGGA	2997
Db	2284	TCCGGGCTTTTCGACCTGCACCAATGGTGGACAAAGGCGAGAGCTGTTCACCGGGTGGT	2343

QY	2998	GAATTTCAATCCAGATAGCAGTACGATTAACCTATATGCGCGTCTCTAAAGGATAGATTTACA	305
Db	2344	GCCCATCT--GGTCGAGCTGGAGCGACGTAACAGCG-----CACAGTTTCAGCGTGTCC	2398
QY	3058	ATATCGCGAGACAACGCCAAGACACACATTTGTTCTCTGCAATGGACAGCCTTGACAGCCCGAA	3117
Db	2399	GGCAGGGCGAGGGCGATGCCACTACGGCAAGCTGACCCCTGAAGTTTCATCTGCACCACC	2458
QY	3118	GACACCGGGGTCTATTTTGTGCAAGCCTTTACTTTCGGCTTCCCTTGGTTTGTCTATTGG	3177
Db	2459	GGCAAGCTGCCGTGCCCTGGCCCACTCGTGACCACTTTCACCTACGGCGTGCAGTGC	2518
QY	3178	GGCCAAAGGACCCCGGTCAACCGTCTCCAGTGTCTAAGCCCAACACGACGCGCAGCGCCGCGA	3237
Db	2519	TTACGCGCTACCCCGACCAATGAAGCAGCAGACTTCTTCAAGTCCGCAATGCCCGAA	2578
QY	3238	CCACCAACACGGGGCCCAACATCGCGTTCGAGCCCTGTCTCCCTGCGCCACAGAGGCGCT	3297
Db	2579	GGCTACGTCGAGGAGCGCAACATCTTCTCAAGGACGAG-----GCAACTCAAGACC	2632
QY	3298	CGGCCAGCGCGGGGGCGCAGTGCAACAGAGGGGGCTGGACTTCGCCCTTGGATCCCAAA	3357
Db	2633	CGCGCGAGGTGAAGTTTCGAGGGCGACACCTGGTGNACCGCATCGAGCTGAAGGGCATC	2692
QY	3358	CTCTGTACTCTGTGTGAATGAAATCTCTTCATCTATGGTGTCAATTCATCTGCTGTGTTTC	3417
Db	2693	GACTTCAAGGAGACGGCAACTCTGGGGCAACAAGCTGGAGTCAACTACAACAGCCAC	2752
QY	3418	CTGAGAGTAAGTTACGACGAGGCGCAGAGCCCGCCGCTACACGAGGCGCCAGAACCCAG	3477
Db	2753	AACGCTTATATCATGCGCGCAACAGCAGAGAAGACGGGCATCAAGGTGAATTTCAAGATCCGC	2812
QY	3478	CTCTATAACGAGCTCAATCTAGGACGAAGAGAGAGTACCATGTTTGTGGACAAGAGACGT	3537
Db	2813	CACAAATCATGAGAGCGCAGCTGCAGCTGCCGACCACTACACGACGAACAACCCCATC	2872
QY	3538	GGCGGACCCCTGAGATGGGGGAAAACCGAGAAAGGAAGAACCCCTCAGGAAGCCCTGTATC	3597
Db	2873	GGCGACGGCCCGTGTCTGTCTGCCGACAACCACTACTCTGAGCACCACTAGTCCGCCCTGAGC	2932
QY	3598	AATGAATCTGAGAAAGATAGATGGCGGAGCCTACAGTGCAGATTTGGGATGAAGGGCGAG	3657
Db	2933	AAAGACCCCAACGAGAAACGCGATCAATGGTCTCTGCTGAGTTTCGTGACCCGCGCGGG	2992
QY	3658	CGCGGAGGGCGCAAGGGGCAAGTGGCTTTTACAGGGTCTCAGTACAGCCACAAGGAC	3717
Db	2993	ATCACTCTCGCATGAGCA-----GCTGTACAAGTAAATGAATTAATTAAG	3038
QY	3718	ACCTACGACCCCTTCAATGACGAGCGCTGCCCTCTGCTAACTCGACGCGCGCGGGAT	3777
Db	3039	AATTGCGCGCGTGCAGCTTGCAGCC-----AAGCTTATCGATAAAATAAAAGATTTTATT	3094
QY	3778	CCGATTAGTTCGAATTTGTTTAAAGACAGGATATCAGTGGTTCAGGCTCTAGTTTGTACTC	3837
Db	3095	TAGTCTCCAGAAAAGGGGGGAATGAAGAACCACCTGTGATGTTTGGCAAGAAATTCGTT	3154
QY	3838	AACAATATCACAGCTGAACCTATAGATACGAGCCATAGATAAAATAAAGATTTTAT	3897
Db	3155	TAAACGGGCTCGAGTGTCTCATACAGACTTATAGATTTCCCAATTCCAAGACATTTTC	3214
QY	3898	TTAGTCTCCAGAAAAGGGGGGAATGAAGACCCCACTCTAGTGTTTGGCAAGCTAGCTT	3957
Db	3215	ACGTTTATGGTATTTCCAGAACACATACGGCATGCAAAATATTGGGATCCCGTAGCTT	3274
QY	3958	AGTTAAGCCCATTTTGAAGCGATGGNAAAATACATATCTGAGAAATAGAGAAGTTTCAGATC	4017
Db	3275	AAGTTAAGCCCATTTTGAAGCGCATGGAAAAATACATTAATCTGAGAAATAGAGAAGTTTCAGATC	3334
QY	4018	AAGTTTAGGAACAGAGAGACAGAGATATGGGCGCAACAGGATATCTGTGGTAAAGCAGT	4077
Db	3335	AAGTTTAGGAACAGAGAGACAGC-----GCTGTAGGAGAGCAGC-----	3357
QY	4078	TCCTGCCCGCTCAGGGCCGAAGAACAGTGTGGAAACAGGAGAAATATGGGCCAAACAGGATAT	4137

QY 6296 GGAGGGCTTACCATCTGGCCCGAGTGTCAATGATACCGCGAGACCCACGCTCAACCGGC 6355
 Db 5182 GGAGGGCTTACCATCTGGCCCGAGTGTCAATGATACCGCGAGACCCACGCTCAACCGGC 5241
 QY 6356 TCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTTCG 6415
 Db 5242 TCCAGATTTATCAGCAATAAACCCAGCAGCCGGAAGGCGGAGCGAGAGTGGTCTTCG 5301
 QY 6416 AACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC 6475
 Db 5302 AACTTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTC 5361
 QY 6476 GCCAGTTAATAGTTTGGCAACGTTGTTGCCAATGCTACAGC-TGCTGGTGTCAAGTTC 6534
 Db 5362 GCCAGTTAATAGTTTGGCAACGTTGTTGCCAATGCTACAGC-TGCTGGTGTCAAGTTC 5421
 QY 6535 GTCTGTTGTTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATC 6594
 Db 5422 GTCTGTTGTTATGGCTTCAATCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATC 5481
 QY 6595 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCGATCGTCTCCGATCGTTGTGAGAAGTAA 6654
 Db 5482 CCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCCGATCGTCTCCGATCGTTGTGAGAAGTAA 5541
 QY 6655 GTTGGCCGAGTGTATCACTCATGTTATGTCAGCAGTGCATATTCCTTACTGTCTAT 6714
 Db 5542 GTTGGCCGAGTGTATCACTCATGTTATGTCAGCAGTGCATATTCCTTACTGTCTAT 5601
 QY 6715 GCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTCGAATA 6774
 Db 5602 GCCATCCGTAAGATGCTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATCTCGAATA 5661
 QY 6775 GTGTATGCGGCGACGAGTTGCTTTGCGCGGCTCAATACGGGATAATACCGCGCCACA 6834
 Db 5662 GTGTATGCGGCGACGAGTTGCTTTGCGCGGCTCAATACGGGATAATACCGCGCCACA 5721
 QY 6835 TAGCAGAACTTTAAAGTCTCATCATGTTGGAACGTTCTTGGGGCGAAGTCTCTCAG 6894
 Db 5722 TAGCAGAACTTTAAAGTCTCATCATGTTGGAACGTTCTTGGGGCGAAGTCTCTCAG 5781
 QY 6895 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCAACTGATCTTC 6954
 Db 5782 GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCAACTGATCTTC 5841
 QY 6955 AGCATCTTTTATCAGCGGTTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGC 7014
 Db 5842 AGCATCTTTTATCAGCGGTTTCTGGGTGAGCAAAAAAGGAGGCAAAATGCCGC 5901
 QY 7015 AAAAAAGGAATAAGGGCGACAGGAAATGTTGAATCTCATCTCTCTTTTCAATA 7074
 Db 5902 AAAAAAGGAATAAGGGCGACAGGAAATGTTGAATCTCATCTCTCTTTTCAATA 5961
 QY 7075 TTATTGAAGCATTTATCAGGTTATTGTTCTCATGAGCGATACATATTTGAATGATTTA 7134
 Db 5962 TTATTGAAGCATTTATCAGGTTATTGTTCTCATGAGCGATACATATTTGAATGATTTA 6021
 QY 7135 GAAAAATAAACAAATAGGGGTTCCGCGCATTTTCCCGGAAAGTGCCACTGAGTCTTA 7194
 Db 6022 GAAAAATAAACAAATAGGGGTTCCGCGCATTTTCCCGGAAAGTGCCACTGAGTCTTA 6081
 QY 7195 AGAACCATTATATCATGACATTAACCTATAAATAAGCGGTATACAGAGCCCTTTCG 7254
 Db 6082 AGAACCATTATATCATGACATTAACCTATAAATAAGCGGTATACAGAGCCCTTTCG 6141
 QY 7255 TCTCGCGCTTTTCGGGTGATGACGGTGAACCTCTTGACACATGACGCTCCCGGAGCGT 7314
 Db 6142 TCTCGCGCTTTTCGGGTGATGACGGTGAACCTCTTGACACATGACGCTCCCGGAGCGT 6201
 QY 7315 CACAGCTTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAAGGCGCGTCAAGCGG 7374
 Db 6202 CACAGCTTCTGTAAAGCGGATGCCGGGAGCAGACAAGCCCGTCAAGGCGCGTCAAGCGG 6261

QY 7375 TGTTCGCGGCTGTCGGGCTGCTTAACTATTCGGCATCAGAGCAGATTGTACTGAGACT 7434
 Db 6262 TGTTCGCGGCTGTCGGGCTGCTTAACTATTCGGCATCAGAGCAGATTGTACTGAGACT 6321
 QY 7435 GCACATATGCGGTGTGAAATACCGACAGATGCTTAAGGAGAAATACCGCATCAGGCG 7494
 Db 6322 GCACATATGCGGTGTGAAATACCGACAGATGCTTAAGGAGAAATACCGCATCAGGCG 6381
 QY 7495 CCAATTGCGCANTTCAGCTGCGCAACTGTTGGGAAGGCGATCGGTGCGGGCTCTTCGCT 7554
 Db 6382 CCAATTGCGCANTTCAGCTGCGCAACTGTTGGGAAGGCGATCGGTGCGGGCTCTTCGCT 6441
 QY 7555 ATTACGCGAGCTGCGGCAAGGGGATGCTCTCAAGGCGATTAAAGTTGGGTAAACGCGAG 7614
 Db 6442 ATTACGCGAGCTGCGGCAAGGGGATGCTCTCAAGGCGATTAAAGTTGGGTAAACGCGAG 6501
 QY 7615 GTTTTCCAGTCAACGCTTTGTAACCGACGCGCCAGTGCC 7654
 Db 6502 GTTTTCCAGTCAACGCTTTGTAACCGACGCGCCAGTGCC 6541

RESULT 4

US-10-789-938B-3
 ; Sequence 3, Application US/10789938B
 ; Publication No. US20050009180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yang, Lili
 ; APPLICANT: Van Parijs, Luk
 ; APPLICANT: Baltimore, David
 ; TITLE OF INVENTION: METHOD FOR THE GENERATION OF
 ; TITLE OF INVENTION: ANTIGEN-SPECIFIC LYMPHOCYTES
 ; FILE REFERENCE: CALTE.008CP1
 ; CURRENT APPLICATION NUMBER: US/10/789,938B
 ; PRIOR FILING DATE: 2004-02-27
 ; PRIOR APPLICATION NUMBER: 10/317,078
 ; PRIOR FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: 60/394,803
 ; PRIOR FILING DATE: 2002-07-08
 ; PRIOR APPLICATION NUMBER: 60/339,375
 ; PRIOR FILING DATE: 2001-12-10
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 7277
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: This represents a retroviral vector encoding a
 ; OTHER INFORMATION: T-cell receptor that recognizes an epitope of
 ; OTHER INFORMATION: Mart-1.
 US-10-789-938B-3

Query Match 43.9%; Score 3359.8; DB 19; Length 7277;
 Best Local Similarity 69.4%; Pred. No. 0;
 Matches 4952; Conservative 0; Mismatches 2022; Indels 159; Gaps 22;
 QY 552 TGAAGACCCCACTGTAGGTTTGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGCAT 611
 Db 1 TGAAGACCCCACTGTAGGTTTGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGCAT 60
 QY 612 GGAATAACATCACTAGAGTAAGTTCAGATCAAGGTTAGGAAACAGAGACAGCA 671
 Db 61 GGAATAACATCACTAGAGTAAGTTCAGATCAAGGTTAGGAAACAGAGACAGCA 119
 QY 672 GAATATGGGCCAAACAGGATATCTGTGTAGTTCCTGCCCGCTCAGGCCCAAGAA 731
 Db 120 ----- 119
 QY 732 CAGTTGGAACAGGAGAAATATGGCCAAACAGAGATCTGTGTAGTTCCTGCCCGC 791
 Db 120 -----AGAATATGGGCCAAACAGGATATCTGTGTAGTTCCTGCCCGC 166
 QY 792 GCTCAGGCCCAAGAACAGATGTTGCCCGCTCCCGCTCAGCAGATTTTCTAGAGA 851

[illegible]

Db	1222	GTCCCCACATCGTGACCTGGGAAGCCCTTGGCTTTTGAACCCCCCTCCCTGGGTCAAGCCC	1281
Qy	1920	TTTGTACACCTAAGCTCGCGCTCTCTTCTTCATCCGCGCCGCTCTCCCCCTTGAA	1979
Db	1282	TTTGTACACCTAAGCTCGCGCTCTCTTCTTCATCCGCGCCGCTCTCCCCCTTGAA	1341
Qy	1980	CCTCCTCGTTGGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTCTCTAGGC	2039
Db	1342	CCTCCTCGTTGGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACTCCTCTCTAGGC	1401
Qy	2040	GCCCCCATATGGCCATATAGATCTTATATGGGGCACCCGCGCCCTTGTAAACTCCCT	2099
Db	1402	GCCGAGATCTCTCGAGT-----TAAAGAAACGGATCCATGTTGCTTGAACATTTATTA	1455
Qy	2100	GACCCTGACATGACAAGAGTTACTTAACAGAGCCCTCTCTCCAAGCTCACTTACAGGCTTCT	2159
Db	1456	ATAATCTTGGATGCGAGCTGCATATGGGTGAGTGTCAACAGCTGAATCAGATCCTCAA	1515
Qy	2160	ACTTAGTCCAGCACGAAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAACAACTGGACC	2219
Db	1516	TCTATGTTTATCCAGGAAGGAGAAGATGCTCCATGAATGCATCTTCTCAAGCATATTT	1575
Qy	2220	GACCGTGG-TACCTCACCCCTTACCGAGTCGGGACACAGTGTGGGTCCGCCGACACAG	2278
Db	1576	AACACCTGGCTATGTGTAACAAGCAGGACCCCTGGGGAGGTCCTGTCTCTTTGATAGCCTTA	1635
Qy	2279	ACTAAGAACCTAGAACCTCGCTGGAAGGACCTTACACAGTCTGCTGACCAACCCCAACC	2338
Db	1636	TATAGGCTGGTGAATTGACCTCAATGGGAAGACTGACTGCTCAGTTTGGTATTAACCGA	1695
Qy	2339	GCCCTCAAAGTAGACGGCATCGCAGCTTGGATATACCGCCCAACGCTGAAGGCTGCCGAC	2398
Db	1696	AAGSACAGCTTCTGTAATATCTCAGCATCCATACCTAGTGTAGGCAT-----CTAC	1749
Qy	2399	CCGGGGGTGACCATCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTTGG	2458
Db	1750	TTCTGTCTGTGGGACCGGTAAACCAAGTTCTATTTTGGGACAGGGAACAAGTTTGACGGTC	1809
Qy	2459	TAGCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCCTGA	2518
Db	1810	ATTCCAAATATCAGAAACCTTGACCTTCCCGTGTACAGCTGAGAGACTCTAAATCCAGT	1869
Qy	2519	GCGCCAGCGTGGGTGACAGAGTGACCATCACCTGTGAAGGCCAGTCAAGGTCTGGTACTT	2578
Db	1870	GA-----CAAGTCTGTCTGCTATTACCGATTTTGATTTCTCAACAATGTGTCAAAA	1924
Qy	2579	CTGTAGCTTGGTATCCAGCAGAACGAGGTGAAGGCTCCAAAGTGTGATCTACTGGACAT	2638
Db	1925	GTAAGGATTTGTATGTATATCACAGACAAAACCTGTGCTAGACATGAGGTCTATGGACT	1984
Qy	2639	CCACCGGCACACTGGTGTGCCAAGCAGATTCACGCGTAGCGGTAGCGGTACCGACTTCA	2698
Db	1985	TCAAGACAAAGTGTGTGGCTTGAGAGCAACAAATCTGACTTTGCAATGTGCAAAACGCT	2044
Qy	2699	CCTTCCACATCAGCAGCCTCCAGCCAGAGGACATCGCCACTACTACTGCCAGCAATATA	2758
Db	2045	TCAACAACAGCATTAATTCAGAAAGACACCTTCTTCCCGAGCCAGAAAGTTCTGTGATG	2104
Qy	2759	GCCTCTATTCGGTGTTCGGCCAAAGGACCAAGGTGGAAATCAAACGAGGTGGCTCAGGAT	2818
Db	2105	TCAAGCTGTGCGAGAAAAGCTTTTGAACAGATACGAAACCTTAACTTTTCAAAAACCTGT	2164
Qy	2819	CGGGTGGATCCGG-CTCTGTGGCTCAGATCGGAGGTCCAACTGGTGGAGAGCGGTGG	2876
Db	2165	TGANTGGTTCCGAATCCTCCTCGAAGGTGGCGGGTTTAATCTGCTCATGACGCTGC	2224
Qy	2877	AGGTGTTGTGCAACCTGGCGGTCTCTGGCGCTGTCTGCTCGGCATCTGGCTTCGATTT	2936
Db	2225	GGCTGTGGTCCAGCTGAGAAATTCGGCTGAGGCTGTGGTCCAGTTTGAAGCCCTCTCCCTC	2284
Qy	2937	CACACATATTGATGAGTTGGGTGAGAC-----AGGCACTGGAAAAGGTCTTTGAGTGA	2992
Db	2285	CCCCCCCCCTAAAGTTACTGGCCGGAAGCGCTTGGAAATAGGCCGGGTGTCGTTTGTCTA	2344

QY 2993 TTGGAGAAATTCATCCAGATAGCAGTACGATTAACATATGCGCGTCTCTAAAGGATAGAT 3052
 Db |||||
 QY 2345 TATGTTATTTCCACCATATGCGTCTTTTGCAATGTAGGGCCCGGAACCTGGCCC 2404
 Db |||||
 QY 3053 TTACAATATCGGAGACAAACGCAAGAAACACATTTGTTCTGCAATGGAACG-CTGAGA 3111
 Db |||||
 QY 2405 TGTCTTTCTTGACGCAATCTCTAGGGTCTTTCCCTCTCGCCCAAGGAATGCAAGGTCT 2464
 Db |||||
 QY 3112 CCGGAGACACGGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGGTTTGTCT 3171
 Db |||||
 QY 2465 GTTGAATGTCGTAAGGAAGCAGTCTCTGGAAGCTTCTTGAAGCAAAACAACGCTCTGT 2524
 Db |||||
 QY 3172 TATTGGGGCAAGGACCCCGGTACCGTCTCCAGTGTAAAGCCACCAAGCGCCAGCG 3231
 Db |||||
 QY 2525 AGGACCTTTTGCAGCAGCGAACCCTCCACTGCGCAGAGTGGCTCTGCGGCCAAAA 2584
 Db |||||
 QY 3232 CCGGACCAACCAACCGCGGCCACCATCGGTGCGAGCCCTGTCTCCCTGGCGCCAGAG 3291
 Db |||||
 QY 2585 GCCACGTGTATAAGATACACCTGCAAGGCGGCACAAACCCAGTGCACGTTGTGAGTTG 2644
 Db |||||
 QY 3292 GCGGCTGCGCAGCGCGGGGCGAGTGCAACAGAGGGGTGCACTTCGCCCTGGAT 3351
 Db |||||
 QY 2645 GATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCGTATTTCAACAAGGGGTGAAGGA 2704
 Db |||||
 QY 3352 CCCAAACTCTGTACTCTGTGATGGAATCCTCTTCTCATCTATGGTGTCAATTCATCTGCC 3411
 Db |||||
 QY 2705 TGGCCAGAGGTACCCATTTGATGGGATCTGATCTGGGGCTCGGTGCAATGCTTTTAC 2764
 Db |||||
 QY 3412 TTGTTCTGAGAGTGAAGTTACGAGAGCGCAGAGCCCGCGGTACAGCAGGCGCCAG 3471
 Db |||||
 QY 2765 ATGTGTTT--AGTCAGGTAAATAACGCTTAGGCCCCCGAACCAACCGGGACGTGTT 2821
 Db |||||
 QY 3472 AACAGCTCTATAACAGCTCAATCTAGAACAGAGAGGATGACATGTTTGTGACAAAG 3531
 Db |||||
 QY 2822 TTCTTTTGAATAACAGATGATATATGCCCACCAATGCGGCACAAAGTTGTTCTTCTA 2881
 Db |||||
 QY 3532 AGACGTGGCGGACCTCAGATGGGGGAAGCCGGAAGAGNAGNAGCCCTCAGGAAGC 3591
 Db |||||
 QY 2882 TGTGGCCCTTGTCTCTGTGACAGGACACATGGATGTGGAATCCTCCAGAGCCCAAG 2941
 Db |||||
 QY 3592 CTGTACAATGAATGCGAGAAAGATAGA-TGGCGGAGGCTACAGTGAGATTGGGATGAA 3650
 Db |||||
 QY 2942 ACACAGGTCACAGACAGGACACCACTGACTCTGAGATGTCAACAGACTGAGAACCA 3001
 Db |||||
 QY 3651 AGCGAGCCCGGAGGGGCAAGGGCAGATGCGCTTTACAGGGTCTCAGTACAGCCAC 3710
 Db |||||
 QY 3002 CCGCTATATGTACTGTATCGACAAGACCCGGGGCATGGGCTGAGGCTGATCCATTACTC 3061
 Db |||||
 QY 3711 CAAGGACCTAGCAGCCCTTACATGCGAGCCCTGCCCCCTCGCTAACTCGACCGGC 3770
 Db |||||
 QY 3062 ATATGTTGTTAAGATAC---TGACAAGGAAGTCTCAGATGGCTATAGTGTCTCTAG 3118
 Db |||||
 QY 3771 CCGGATCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTT 3830
 Db |||||
 QY 3119 ATCAAGACAGAGGATTTCTCTCTACTCTGAGTCCGTACAGCTCCCAACATCTGT 3178
 Db |||||
 QY 3831 TTGACTCAACAATACA-CCAGCTGAAGCTTATAGATGACGACCATAGATATAATAAAA 3889
 Db |||||
 QY 3179 GTACTTCTGTGCCATCAGTGAGTAGGGTTGGGAGCCCGCAGCATTTTGGTGATGGGAC 3238
 Db |||||
 QY 3890 GATTTTATTAGTCTCCAGAAAAGGGGGGAATGAAGACCCCACTGTAGTGTGGCA 3949
 Db |||||
 QY 3239 TCGACTCTCCATCTAGAGGACCTGAACAAGGTGTTCCACCGGAGGTGCTGTGTTGA 3298
 Db |||||
 QY 3950 GCTAGCTTAAGTAAACCCATTTTGAAGGCATGGAATAATACATACTAGAAATAGAGAAG 4009
 Db |||||
 QY 3299 GCCATCAGAGCAGAGATCTCCACACCCAAAGGCCACACTGGTGTGCTGGCCACAGG 3358
 Db |||||
 QY 4010 TTCAGATCAAGGTTAGGAACAGAGAGACAGCAAGATATGGGCCAAACAGGATATCTGTGG 4069
 Db |||||
 QY 3359 CTTCTTCCCTGACACGCTGAGCTGAGTGGTGGTGAATGGGAAGGAGTGCACAGTGG 3418
 Db |||||

QY 4070 TAAAGCAGTTTCTGCCCCGCTCAGGGCCCAAGAACAGTTGAAACAGAGAAATATGGGCCAAA 4129
 Db |||||
 QY 3419 -GGTCAGCAGCGACCCGAGCCCTCAAGGAGCAGCCCGCTCAATGATCCAGATACT 3477
 Db |||||
 QY 4130 CAGGATATCTGTGTAAAGATTTCTGCCCCGCTCAGGGCCCAAGAACAGATGGTCCCCA 4189
 Db |||||
 QY 3478 GCTGAGCAGCGCCGCTGAGGGTCTCGGCCACCTTTCTGGCAGAACCCCGCAACCACTTCC 3537
 Db |||||
 QY 4190 GATCGGTCCCGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCCAA 4249
 Db |||||
 QY 3538 GCTGTCAATCCAGTTCTACGGGCTCTCGGAGAAATGACAGTGGACCCAGGATAGGGCCA 3597
 Db |||||
 QY 4250 GGACCTGAAATGACCTGTGCTTATTTGAACTAAACCAATCAGTTCTGCTTCTG 4309
 Db |||||
 QY 3598 AACCCGTACCCAGATGTCAGCGCGAGGCTTGGGTAGACAGACTGTGCTTTTACCT 3657
 Db |||||
 QY 4310 TTTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCTGGCGGC 4369
 Db |||||
 QY 3658 CGGTGTCTTACCCAGCAAGGGTCTGTCTGCCACCATCTCTATGAGATCTCTGTAGGA 3717
 Db |||||
 QY 4370 CAGTCTCGATAGACTGGTCCCGGGTACCGGTGTTCTCAATAAACCCCTTTGCACT 4429
 Db |||||
 QY 3718 AGGCCACCTGTATGCTGTGTCAGCGCCCTTGTGTGATGCCATGGTCAAGAGAA 3777
 Db |||||
 QY 4430 TGCATCCGACTCGTGTCTGCTGCTTCTTGGAGGGTCTCTCTCAGTGATGACTACCC 4489
 Db |||||
 QY 3778 AGGATTTCTGAGTCACTGCGAGCCAAAGCTTATCATATAAATAAAGATTTTATTAGTC 3837
 Db |||||
 QY 4490 GTACGCGGGTCTTTCAGTTCTCCCACTACACAGGTCTCACTAACATCTCTGATGTC 4549
 Db |||||
 QY 3838 TCCGAAAAAAGGGGGGAATGAAAGACCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAA 3897
 Db |||||
 QY 4550 CGCAGGACTCGTCCAGCCCGTCTTGTATATAAATAAAGTCAAGAACAGTGTCCCTT 4609
 Db |||||
 QY 3898 CGCCATTTTGCAGGCATGGAATAACATACTGAGATAGAGAGTTGAGTCAAGGTT 3957
 Db |||||
 QY 4610 CAAGCAGACTACATCTGACTCTCGGCTTTATAAAGAAATGTTCAAGGGC---TCTGT 4665
 Db |||||
 QY 3958 AGGAACAGAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGTAGCAGTTCCTGC 4017
 Db |||||
 QY 4666 GGACTATCTGCCACACGACTTTTAAAGATTTTATGCTCTCTGATGAGGATTTAGTCA 4725
 Db |||||
 QY 4018 CCGGCTCAGGGCCAAAGAACAGATGGTCCACAGATGCGGTCCCGCCCTCAGCAGTTCTA 4077
 Db |||||
 QY 4726 ATCTATCTCTGTTATTTTGTGGCTTCTCGTATTTTAAATTTCTAGTTGCACTCCCT 4785
 Db |||||
 QY 4078 GAGAACCATCAGATGTTTCCAGGGTGGCCCAAGGACCTGAAAATGACCCCTGTGCTTAT 4137
 Db |||||
 QY 4786 TCCTGAGAGCAGCGGCTTTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAG 4845
 Db |||||
 QY 4138 TGAACATAACCAATCAGTTCGC-----TTCGCTTCTGTGCGCGCTTCTGCTCCCG 4191
 Db |||||
 QY 4846 GTTGTCTGGGCTCAGTGTGAGATTTTGGCATAAAGGGGTCCTGCCCTGTGTACAGAC 4905
 Db |||||
 QY 4192 AGCTCAATAAAGAGCCCAACCCCTCACTCGGCGCCAGTCTCCGATAGACTGGGT 4251
 Db |||||
 QY 4906 AGATCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGT-----CGGGGCTCTGATCTCA 4962
 Db |||||
 QY 4252 CGCCCGGGTACCCGCTGATCCAAATAAACCCCTCTTCAGTTGATCCGACTTGTGGTCTCG 4311
 Db |||||
 QY 4963 GGGCATCTTTCCTAGAGATCTCTACCGCGAGCATCTGTCGGCGGTACCGAGTCCA 5022
 Db |||||
 QY 4312 CTGTTCTTGGAGGGTCTCTCTGAGTGTGACTACCGGTACGCGGGGTCTTTCAGT 4371
 Db |||||
 QY 5023 ATTCGTAATCATGTCATAGTCTTCTGTTGTGAAATTTGTTTATCCGCTCACAATTTCCAC 5082
 Db |||||
 QY 4372 ATTCGTAATCATGTCATAGTCTTCTGTTGTGAAATTTGTTTATCCGCTCACAATTTCCAC 4431
 Db |||||
 QY 5083 ACAACATACGAGCCGGAAGCATATAAGTGTAAAGCTGGGGTCCCTAATGAGTGAAGTAAAC 5142
 Db |||||
 QY 4432 ACAACATACGAGCCGGAAGCATATAAGTGTAAAGCTGGGGTCCCTAATGAGTGAAGTAAAC 4491
 Db |||||
 QY 5143 TCACATTAATTGCGTTGCGTCACTGCGCGCTTTTCCAGTTCGGGAAACCTGTGTCGACAGC 5202
 Db |||||

Db 4492 TCACATTAATTGCGTTGGCTCACTGCCCCGTTTCCAGTCGGGAACCTGTCGGCAGC 4551
Qy 5203 TGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTTCCG 5262
Db 4552 TGCATTAATGAATCGGCCAACCGCGGGAGAGCGGTTTCGCTATTGGCGCTCTTTCCG 4611
Qy 5263 CTTCTCGCTCACTGACTCGCTGCGCTCGCTGCGCTGCGCTGCGCGGAGCGGTATCAGGTC 5322
Db 4612 CTTCTCGCTCACTGACTCGCTGCGCTCGCTGCGCTGCGCTGCGCGGAGCGGTATCAGGTC 4671
Qy 5323 ACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGCGAGGAAGAACATGT 5382
Db 4672 ACTCAAGGCGGTAAATACGGTTATCCACAGATCAGGGGATAACGCGAGGAAGAACATGT 4731
Qy 5383 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCTGATAAAAGCGCGCTGCTGCGCTTTTCC 5442
Db 4732 GAGCAAAAGGCCAGCAAAAGGCCAGGAACCTGATAAAAGCGCGCTGCTGCGCTTTTCC 4791
Qy 5443 ATAGGCTCCGCCCTCGACAGGATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 5502
Db 4792 ATAGGCTCCGCCCTCGACAGGATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAA 4851
Qy 5503 ACCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 5562
Db 4852 ACCGACAGGACTATAAGATACCAAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTC 4911
Qy 5563 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 5622
Db 4912 CTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGG 4971
Qy 5623 CGCTTTCTCATAGCTCAGGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGC 5682
Db 4972 CGCTTTCTCATAGCTCAGGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAAGC 5031
Qy 5683 TGGGCTGTGTCACGAACCCCGGTTGAGCCGACCGCTGCGCTTATCCGGTAACATATC 5742
Db 5032 TGGGCTGTGTCACGAACCCCGGTTGAGCCGACCGCTGCGCTTATCCGGTAACATATC 5091
Qy 5743 GTCTTTGAGTCAACCCCGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGTAACA 5802
Db 5092 GTCTTTGAGTCAACCCCGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGTAACA 5151
Qy 5803 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTCTGAAGTGTGGCGCTAACT 5862
Db 5152 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTCTGAAGTGTGGCGCTAACT 5211
Qy 5863 ACGGCTACCTAGAAAGGACGATATTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCG 5922
Db 5212 ACGGCTACCTAGAAAGGACGATATTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCG 5271
Qy 5923 GAAAAAGTTGGTAGCTTTGATCCGGCAACAAACCCAGCTGGTAGCGGTGGTTTTT 5982
Db 5272 GAAAAAGTTGGTAGCTTTGATCCGGCAACAAACCCAGCTGGTAGCGGTGGTTTTT 5331
Qy 5983 TTGTTTGAACGACGAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCT 6042
Db 5332 TTGTTTGAACGACGAGATTACGGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCT 5391
Qy 6043 TTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAAGTTAAGGATTTTGGTCATGA 6102
Db 5392 TTTCTACGGGCTGACGCTCAGTGGAAACGAAAACTCAAGTTAAGGATTTTGGTCATGA 5451
Qy 6103 GATTATCAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAA 6162
Db 5452 GATTATCAAAAGGATCTTACCTAGATCCTTTTAAATTAATAATGAAGTTTAAATCAA 5511
Qy 6163 TCTAAAGTATATAGTAAGTAACTTGGCTGACAGTTACCAATGCTTAATCAGTCAGGCAC 6222
Db 5512 TCTAAAGTATATAGTAAGTAACTTGGCTGACAGTTACCAATGCTTAATCAGTCAGGCAC 5571
Qy 6223 CTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCGCTCGGTAGA 6282

Db 5572 CTATCTCAGCGATCTGTCTATTTCGTTTCATCCATAGTTGCTGACTCCCCGTCGTAGA 5631
Qy 6283 TAACTACGATACCGGAGGGCTTACCATCTGCCCCAGTGTGCTGAATGATACCGGAGACC 6342
Db 5632 TAACTACGATACCGGAGGGCTTACCATCTGCCCCAGTGTGCTGAATGATACCGGAGACC 5691
Qy 6343 CACGCTCACGGCTCCAGATTATACGCAATAAACACGAGCCAGCGGAGGCGCGAGCGCA 6402
Db 5692 CACGCTCACGGCTCCAGATTATACGCAATAAACACGAGCCAGCGGAGGCGCGAGCGCA 5751
Qy 6403 GAAGTGTCTCTGCAACTTTATCCGCTTCATCCAGTCTATTAAATTTGTCGGGAAGCTA 6462
Db 5752 GAAGTGTCTCTGCAACTTTATCCGCTTCATCCAGTCTATTAAATTTGTCGGGAAGCTA 5811
Qy 6463 GAGTAAGTAGTTCCGCAAGTTAATAAGTTTGGCAACGTTGTGGCACTTGTCTACAGCC-TCG 6521
Db 5812 GAGTAAGTAGTTCCGCAAGTTAATAAGTTTGGCAACGTTGTGGCACTTGTCTACAGCCATCG 5871
Qy 6522 TGGTGTCAAGCTCGTGGTTTGGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGGC 6581
Db 5872 TGGTGTCAAGCTCGTGGTTTGGTATGCTTCATTCAGCTCCGTTCCCAACGATCAAGGC 5931
Qy 6582 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTCGGTCTCCGATCG 6641
Db 5932 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTCGGTCTCCGATCG 5991
Qy 6642 TTGTCAAGTAAGTTTGGCGCGAGTGTATCACTCATGTGTTATGCGAGCACTGCATAATT 6701
Db 5992 TTGTCAAGTAAGTTTGGCGCGAGTGTATCACTCATGTGTTATGCGAGCACTGCATAATT 6051
Qy 6702 CTCTTACTGTCAATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGT 6761
Db 6052 CTCTTACTGTCAATCCGTAAGATGCTTTTCTGTGACTGTGTGAGTACTCAACCAAGT 6111
Qy 6762 CATCTCAGAAATAGTGTATGCGGCGACCGAGTGTCTCTGCCCGCGCTCAATACGGGATA 6821
Db 6112 CATCTCAGAAATAGTGTATGCGGCGACCGAGTGTCTCTTGCCCGGCGTCAATACGGGATA 6171
Qy 6822 ATACCGGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGCTTTCTCGGGC 6881
Db 6172 ATACCGGCGCACATAGCAGAACTTTTAAAGTGTCTCATCTTGGAAAAAGCTTTCTCGGGC 6231
Qy 6882 GAAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGATAACCACTCGTGAC 6941
Db 6232 GAAAACTCTCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGATAACCACTCGTGAC 6291
Qy 6942 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAAACGAA 7001
Db 6292 CCAACTGATCTTCAGCATCTTTTACTTTTCCAGCGGTTTCTGGGTGAGCAAAAAACGAA 6351
Qy 7002 GGCAAAATGCCGAAAAAGGGAATAAGGCGCACACGGAATTTGAATCTCATACTCT 7061
Db 6352 GGCAAAATGCCGAAAAAGGGAATAAGGCGCACACGGAATTTGAATCTCATACTCT 6411
Qy 7062 TCCTTTTCAATATTATTGAAGCATTTTATCAGGGTATTGTCTCATGAGCGGATACATAT 7121
Db 6412 TCCTTTTCAATATTATTGAAGCATTTTATCAGGGTATTGTCTCATGAGCGGATACATAT 6471
Qy 7122 TTGAATGTTTGAATAAATAAATAAGGGTTTCGCGCACATTTCCCGGAAAAAGTGC 7181
Db 6472 TTGAATGTTTGAATAAATAAATAAGGGTTTCGCGCACATTTCCCGGAAAAAGTGC 6531
Qy 7182 CACCTGACGCTCAAGAACCATTTATATCATGACATTAACCTATAAAAAATAGGCGTATCA 7241
Db 6532 CACCTGACGCTCAAGAACCATTTATATCATGACATTAACCTATAAAAAATAGGCGTATCA 6591
Qy 7242 CGAGGCCCTTTCTCGCTCGCGGTTTTCGGTGTATGTCGGTGAATAAACCCTCTGACACATGCA 7301
Db 6592 CGAGGCCCTTTCTCGCTCGCGGTTTTCGGTGTATGTCGGTGAATAAACCCTCTGACACATGCA 6651
Qy 7302 TCCCGAGAGCGGTCAAGCTTGTCTGAAGCGGATGCGGGAGCAGACAAGCCCGTCAGG 7361
Db 6652 TCCCGAGAGCGGTCAAGCTTGTCTGAAGCGGATGCGGGAGCAGACAAGCCCGTCAGG 6711

QY 7362 GCGCGTCAGCGGTTGGCGGTTGGCGGTTGGCGGTTAACTATGCGCATCAGACAGA 7421
 DB 6712 GCGCGTCAGCGGTTGGCGGTTGGCGGTTGGCGGTTAACTATGCGCATCAGACAGA 6771
 QY 7422 TTGTACTGAGAGTCACCATATGCGGTGTGAATATCCGCACAGATCGTAAGGAGAAAT 7481
 DB 6772 TTGTACTGAGAGTCACCATATGCGGTGTGAATATCCGCACAGATCGTAAGGAGAAAT 6831
 QY 7482 ACCGATCAGGCGCATTCGCCATTCAGGTCGGCAACTGTGTGGGAAGGCGATCGGTGC 7541
 DB 6832 ACCGATCAGGCGCATTCGCCATTCAGGTCGGCAACTGTGTGGGAAGGCGATCGGTGC 6891
 QY 7542 GGGCCTCTTCGCTATTACGCCAGCTCGCGAAGGCGGATGTGTGCAAGCGCATTAAGTT 7601
 DB 6892 GGGCCTCTTCGCTATTACGCCAGCTCGCGAAGGCGGATGTGTGCAAGCGCATTAAGTT 6951
 QY 7602 GGGTAACGCCAGGTTTTCCAGTCACGAGCTTTGTAATAACGACGGCCAGTGCC 7654
 DB 6952 GGGTAACGCCAGGTTTTCCAGTCACGAGCTTTGTAATAACGACGGCCAGTGCC 7004

RESULT 5

US-10-789-938B-2
 ; Sequence 2, Application US/10789938B
 ; Publication No. US20050009180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yang, Lili
 ; APPLICANT: Van Parijs, Luk
 ; APPLICANT: Baltimore, David
 ; TITLE OF INVENTION: METHOD FOR THE GENERATION OF
 ; FILE REFERENCE: CALTE.008CPI
 ; CURRENT APPLICATION NUMBER: US/10/789,938B
 ; CURRENT FILING DATE: 2004-02-27
 ; PRIOR APPLICATION NUMBER: 10/317,078
 ; PRIOR FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: 60/394,803
 ; PRIOR FILING DATE: 2002-07-08
 ; PRIOR APPLICATION NUMBER: 60/339,375
 ; PRIOR FILING DATE: 2001-12-10
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 7295
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: This represents a retroviral vector encoding a
 ; OTHER INFORMATION: T-cell receptor that recognizes an epitope of
 ; OTHER INFORMATION: gp-100.
 US-10-789-938B-2

Query Match 43.8%; Score 3354.4; DB 19; Length 7295;
 Best Local Similarity 69.6%; Pred. No. 0;
 Matches 4982; Conservative 0; Mismatches 1981; Indels 199; Gaps 24;
 QY 552 TGAAGACCCACCTGTAGTTGGCAAGCTTAAGTAAGCGCCATTTTGGCAAGGCAT 611
 DB 1 TGAAGACCCACCTGTAGTTGGCAAGCTTAAGTAAGCGCCATTTTGGCAAGGCAT 60
 QY 612 GGAATATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGAACAGAGACAGCA 671
 DB 61 GGAATATACATACTGAGATAGAGAGTTTCAGATCAAGTTAGGAACAGAGACAGC- 119
 QY 672 GAATATGGCCAAACAGGATATCTGTGGTAAGCAGTTCTCGCCCGCTCAGGCGCAAGAA 731
 DB 120 ----- 119
 QY 732 CAGTTGGAACAGGAAATATGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCG 791
 DB 120 -----AGATATGGCCAAACAGGATATCTGTGTAAGCAGTTCTCGCCCG 166

QY 792 GCTCAGGGCCAAAGAACAGATGTCCTCCAGATCGGTCCCGCTCAGCAGATTTCTAGAGA 851
 DB 167 GCTCAGGGCCAAAGAACAGATGTCCTCCAGATCGGTCCCGCTCAGCAGATTTCTAGAGA 226
 QY 852 ACCATCAGATGTTTCCAGGGTCCCAAGGACCTG-AAATGACCTGTGCTTATTTGAA 910
 DB 227 ACCATCAGATGTTTCCAGGGTCCCAAGGACCTGAAAATGACCTGTGCTTATTTGAA 286
 QY 911 CTAAACCAATCAGTTCGCTTCTGCTTCTGCTCGCGGCTTCTGCTCCCGAGCTCAATAA 970
 DB 287 CTAAACCAATCAGTTCGCTTCTGCTTCTGCTCGCGGCTTCTGCTCCCGAGCTCAATAA 346
 QY 971 AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTGCGCTCCCGGGTA 1030
 DB 347 AAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCCGATAGACTGCGCTCCCGGGTA 406
 QY 1031 CCGCTATTCCCAATAAAGCCTTCTGCTTCTGCTTCCGATCCGAAATCGTGGACTCGTGTAT 1090
 DB 407 CCGCTATTCCCAATAAAGCCTTCTGCTTCTGCTTCCGATCCGAAATCGTGGACTCGTGTAT 466
 QY 1091 GGGAGGCTCTCCTCAGATTGATGCTGCCACCTCGGGGCTTCTTCAATTCGAGGTTCC 1150
 DB 467 GGGAGGCTCTCCTCAGATTGATGCTGCCACCTCGGGGCTTCTTCAATTCGAGGTTCC 526
 QY 1151 ACCGAGATTGGAGACCCCTGCGCAGGAGACACACGACCCCGCGGGAGGTAAAGCTGG 1210
 DB 527 ACCGAGATTGGAGACCCCTGCGTAGGAGACACACGACCCCGCGGGAGGTAAAGCTGG 586
 QY 1211 CCAGCAACTTATCTGTGTCTGTC-----CGATTGTCTAGTGTCTATGACTGATTTTA 1262
 DB 587 CCAGCGGTGTTCTGCTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
 QY 1263 TCGCGCTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 DB 647 TCGCGCTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
 QY 1323 CTGACGAGTTTCGGAACACCCCGCGCACCTCGGAGAGCTGCCAGGAGCTTCGCGGGCC 1382
 DB 707 CTGACGAGTTTCGGAACACCCCGCGCACCTCGGAGAGCTGCCAGGAGCTTCGCGGGCC 766
 QY 1383 GTTTTTGTGGCCGACCTGAGTCTTAAATCCCGATCGTTTATAGGACTCTTTGGTGACCC 1442
 DB 767 GTTTTTGTGGCCGACCTGAGGAGGAGTCTGATGGAATCCGACCC----- 814
 QY 1443 CCCITAGAGGAGGATATGTTGTTCTGCTAGAGAGAGAGAACTAAACAGTTCCCGCT 1502
 DB 815 -----CGTCAGGATATGTTGTTCTGCTAGAGAGAGAGAACTAAACAGTTCCCGCT 867
 QY 1503 CCGTCTGAATTTTCTGCTTTCGGTTTGGACCCGAGCCGCGCGCTCTGCT---CTGC 1559
 DB 868 CCGTCTGAATTTTCTGCTTTCGGTTTGGACCCGAGCCGCGCGCTCTGCTGCTGCTGCTGCT 927
 QY 1560 TGCAGCATGTTTCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
 DB 928 TGCAGCATGTTTCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987
 QY 1620 GSCCGGGGTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTGCTGCTGCTGCTGCTGCT 1679
 DB 988 GSCC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTGCTGCTGCTGCTGCTGCT 1041
 QY 1680 GAGCGGATCGCTCACAACCCAGTCGCTAGATGTCAGAAGAGAGACGTTGGTTACCTTCTGCT 1739
 DB 1042 GAGCGGATCGCTCACAACCCAGTCGCTAGATGTCAGAAGAGAGACGTTGGTTACCTTCTGCT 1101
 QY 1740 TCTCAGAAATGGCCAAACCTTTAAACGTCGGAATGGCCCGGAGACGGCACCTTTAAACCGAGAC 1799
 DB 1102 TCTCAGAAATGGCCAAACCTTTAAACGTCGGAATGGCCCGGAGACGGCACCTTTAAACCGAGAC 1161
 QY 1800 CTCATCACCCAGGTTAAGATCAAGTCTTTTCACTCGGCGCGCATGGACACCCAGACCCAG 1859
 DB 1162 CTCATCACCCAGGTTAAGATCAAGTCTTTTCACTCGGCGCGCATGGACACCCAGACCCAG 1221
 QY 1860 GTCCCTCATCATCGTGACCTCGGAAAGCCTTGGCTTTTGTAGACCCCTTCCCTCGGCTCAAGCCC 1919

Db 1222 GTCCCTACATCGTACCTGGAGGCCCTGGCTTTTGACCCCTCCCTGGTCAAGCCC 1281
Qy 1920 TTTGTACACCTTAAGCCTCCGCTCCTCTTCTCCTCATCGCCCGCTCTCTCCCTTGAA 1979
Db 1282 TTTGTACACCTTAAGCCTCCGCTCCTCTTCTCCTCATCGCCCGCTCTCTCCCTTGAA 1341
Qy 1980 CCTCCTGTTGACCCCGCTCGATCTCCTCTTTATACAGCCCTCACTCTCTCTAGGC 2039
Db 1342 CCTCCTGTTGACCCCGCTCGATCTCCTCTTTATACAGCCCTCACTCTCTCTAGGC 1401
Qy 2040 GCCCCATAT-----GGCCATATGATCTTATATGGGGCAACCCCGCCCTTGTAACT 2094
Db 1402 GCGGATCTCTCGAGGTTAAGAAACCGATCCATGGTGAAGATCCGGCAATTTTGTG 1461
Qy 2095 TCCCTGACCCCTGACATGATGAAGAGTTACTACAGCCCTCTCTCAAGCTCACTTACAGG 2154
Db 1462 GCTATTTTGTGCTTCAGCTAAGCTGTGAAGTGCCTGCGCAAAATGAAGTGGACAGAT 1521
Qy 2155 CTT-----CTACTTAGTCAGACGAAGTCTGGAGACCTCTGGGGGACCTACCAAGAA 2209
Db 1522 CCTCAGAACCTGACTGCCAGGAAGAGAAATTTATCAATCAACTCACTGAGTTACTCGGTA 1581
Qy 2210 CACTGGACCGGTGTTACCTCACTTACCGAGTCCGGGACACAGTGTGGTCCGC 2269
Db 1582 GGAATAAGTGCTTACACTGGCTGCAACAGCATCCAGGAGGAGATTTGTTCTCTGTTT 1641
Qy 2270 CGACACACAGATTAAGAACTCGCTGGAAGGACCTTACACAGTCTCTGCTGACC 2329
Db 1642 ATGCTGAGCTCAGGGAAGAGAGATGAAGATTAATGGCAATAAATACAGGAA 1701
Qy 2330 ACCCCACCGCCCTCAAAAGTAGAGGCAATCGAGCTTGGATACAGCGCCGCCACGTGAAG 2389
Db 1702 AAGCACAGCTCCCTGCACATCACAGCCTCCCATCCAGAGACTCTGCGCTTACATCTGT 1761
Qy 2390 GCTGCCGACCCGGGGTGGACCATCTCTTAGACTGCCATGGGATGGAGCTGTATCATCC 2449
Db 1762 GCTGCCCTCAATTAATTCAGGAGC--CCAGAGCTGGTATTTGGCCAGGAACCCAGGCTGA 1819
Qy 2450 TCTTCTTGGTAGCAACAGCTACAGGTGTCCTCGACATCCAGCTGACCCAGAGCCCAA 2509
Db 1820 CTATCAACCCAAATATCCAGAACCTTGACCTGCGCTGATACAGCTGAGAGCTCTAAT 1879
Qy 2510 GCAGCTGAGCGCCAGCGTGGGTGACAGAGTGAACCATCACTGTAAGGCCAGTCAGGATG 2569
Db 1880 CCAG-----TGACAAAGTCTGTCTGCCTATTACCGGATTTTGATTTCTCAACAAATG 1930
Qy 2570 TGGGTACTTCTGTAGCTTGGTACAGAGAGAGCCAGGTGAAGCTCCAAAGCTGTATCT 2629
Db 1931 TGTCAAAAGTAAGGATTTCTGATGTGTATATACAGACAAAACTGTGTAGACATGAGGT 1990
Qy 2630 ACTGGACATCCACCCGGCACACTGTGTGTCGAAGCAGATTCAGCGGTAGCGGTAGCGTA 2689
Db 1991 CTATGGACTTCAGAGCAACAGTCTGTGGCTGAGGCAAAATCTGACTTTGATGTG 2050
Qy 2690 CCGACTTCACCTTCAACATCAGACGCTCCAGCCAGAGGACATGCCACCTACTACTGCC 2749
Db 2051 CAAGCGCTTCAACAAACAGCATTTTCAGAAAGACACCTCTCTCCCGAGCCAGAAAGTT 2110
Qy 2750 AGCAATATAGCTCTATCGGTGTTTCGGCCAAAGGGAACCAAGGTGGAATCAAAACAGGTG 2809
Db 2111 CCTGTGATGTCAAGCTGTGAGAAAAGCTTTGAAACAGATACGAACTTAACATTTCAAA 2170
Qy 2810 GCTCAGGATCGGTGGATCCGG--CTCTGTGGCTCAGGATCGGAGGTCCAACTGGTGA 2867
Db 2171 ACCTGTAGTATTTGGGTTCCGAATCTCTCTCTGAAAGTGGCGGGTTTAAATCTGTCA 2230
Qy 2868 GAGCGGTGGAGGTGTGTGCAACCTGGCGCGTCTCTGCGCTGTCTCTCGCATCTGG 2927
Db 2231 TGAAGCTCGGCTGTGGTCTAGCTGAGNAATTCGCTGAGGCTGTGGTCCAGTTGACGCCC 2290
Qy 2928 CTTGATTTTCAACATATGGAATGAGTTGGGTGAGACAGGCACTGGAAAGGTCTTGA 2987

Db 2291 CTCTCCCTCCCTCCCTCCCTTAACGTTACTGCGCGAAGCCG-----CTTGGAAATAAGCCGGT 2346
Qy 2988 GTGATTTGGAGAAATTCATCCAGATAGCAGTACGATTAACATATGCGCGTCTCTTAAAGAA 3047
Db 2347 GTGCGTTGTCTATATGTTATTTTCCACCATATGCGCGTCTTTTGGCAATGTG---AGGG 2403
Qy 3048 TAGATTTACAATATCGGAGAGCAAGCCCAAGAACACATTTGTTCTTGCAAAATGAGACGCT 3107
Db 2404 CCGGAAACCTTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCTCTCGCA 2463
Qy 3108 GAGACCGGAAGACACCGGGTCTATTTTGTGGAAGCCTTTACTTTCGGCTTCCCTGTGTT 3167
Db 2464 AAGAAATGCAAGTCTGTTGAATGTGTGAAGAGCAGATTCCTCTGAAGCTTCTTGA 2523
Qy 3168 TGCTTATTTGGGGCAAGGGACCCCGGTCAACGCTCTCCAGTCTTAAGCCCAACACGACGCC 3227
Db 2524 GACAAACAACTCTGTAGCGACCTTTGCAGG---CAGCGGAACCCCTCCCTGCGACA 2580
Qy 3228 AGCGCGGAGACCAACACACCGGGCCCAACCATCGGTGCGAGCCCTGTCTCTCGGCC 3287
Db 2581 GGTGCTCTGCGGCCAAAGACACGTGTATAGATACACCTTGCAAGCGGCAACAACCC 2640
Qy 3288 AGAGCGGCTCGGCCAGCGCGGGCGCAGTGCACACGAGGGGGCTGGACTTCGCCCT 3347
Db 2641 AGTCCACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAAATGGCTCTCTCAAGCGTAT 2700
Qy 3348 GGATCCCAAACTCTGCTACCTGCTGGATGGAATCTCTTCTATCTATGTGTCATTTCTCAC 3407
Db 2701 TCAACAGGGGCTGAAGATGCCCAAGAGGTACCCCATTTGTATGGATCTGATCTGGGGC 2760
Qy 3408 TGCCTTGTCTGAGAGTGAAGTTACAGAGGCGCAGAGCCCCCGCGTACACGACAGG 3467
Db 2761 CTCGGTGCACATGCTTTACATGTTAGTCGAGGTTAAAAAAGCTGAGGCCCCCGAA 2820
Qy 3468 CCAGAACCGCTCTATAACAGCTCAATCTAGGAGAGAGAGAGTACGATGTTTGA 3527
Db 2821 CCAGGGGACGTGTTTCTTTTGAATAACAGATGAATATGAGGCCCAACCATGGACT 2880
Qy 3528 CAAGAGACGTGGCGGACCTTGAGATGGGGGAAAGCCGAGAGGAAGAACCTTCAGA 3587
Db 2881 CCTGGACTTCTGTG-----TGTTGCCCTTTGATCTCTGTAGGAGCATACAGATGC 2935
Qy 3588 AGGCTGTACAAATGCACTGCAAGAAAGATTAAGATGGCGAGCCCTACAGTGAAGTTGGAT 3647
Db 2936 TGGAGTTATCCAGTCAACCCGCCATGAGTGACAGAGATGGGACAAGAGTGACT---CT 2992
Qy 3648 GAAAGGCGAGCGCGGAGGGCAAGGGCAGATGGCTTTTACAGGCTCTCAGTACAGC 3707
Db 2993 GAGATGTAACCAATTTTCAAGGCCCAACTCCCTTTTCTGGTACAGACAGACATGATCGG 3052
Qy 3708 CACCAAGACACCTTACAGCGCTTTCATGCGGCCCTTCCCTCGCTAACTCGACGC 3767
Db 3053 GGGACTGGAGTTGCTCATTTACTTTTAAACAAAGTTCCGATAGATGATTTCA---GGAT 3109
Qy 3768 GGGCGCGATCCGGATTTAGTCCAAATTTTAAAGACAGGATATCAGTGGTCCAGGCTCTA 3827
Db 3110 GCCCGAGGATCGATTTCTCAGCTAAGTGCCTAATGTCATCAATCTCCTCACTCTGAAGATCCA 3169
Qy 3828 GTTTTGAATCAAAATATCACCAGCTGAGGCTATAGATGAGGACCATAGATAAAATAA 3887
Db 3170 GCCCTCAAGACCCAGGAGCTCAGCTGTGTACTTCTGTGCCAGCAGCCCGGGGCAATGA 3229
Qy 3888 AAGATTTTATTTAGTCTCCAGAAAAAGGGGGAATGAAAGACCCACCTGTAGGTTTGGC 3947
Db 3230 GCAGTTCTTGGGCCAGGACACCGCTTCAACGCTCTAGAGGACCTGAAAAACGTTTCCC 3289
Qy 3948 AAGTAGCTTAAGTAACGCCATTTTGAAGGCATGGAATAATACATACTGAGAAATAGAGA 4007
Db 3290 ACCCGAGGTGCTGTGTTTGAAGCCATCAGAGCAGAGATCTCCACACCCCAAGGCCAC 3349
Qy 4008 AGTTCAGATCAAGGTTTAGGAACAGAGAGACAGCAGAAATATGGGCCCAACAGATATCTGT 4067
Db 3350 ACTGTTATGCTGCGCCACAGGCTTCTACCCCGACCACTGAGAGCTGAGCTGTGGGTGAA 3409

QY 4068 GGTAAGCAGTTCTGTCGCCGCTCAGGGCCAAAGAACAGTTGGAAACAGGAGAAATATGGCC-- 4125
 Db 3410 TGGGAAGGAGGTGCACAGTGGGGTCAGACAGACCAGCCGAGCCCTCAAGGAGCAGCCCGC 3469
 QY 4126 -----CAAACAGAGATATCTGTGTAAAGCATGTTCTGCCCCGGCTCAGGCGCCAAAGAACAGA 4180
 Db 3470 CCTCAATGACTCCAGATACCTGCTGAGCAGCGCCCTGAGGGTCTCGGCCACCTTCTGGCA 3529
 QY 4181 TGGTCCCCAGATGCGGTTCGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGG 4240
 Db 3530 GAACCCCCCAACCACTTCGGCTGTCAAGTCCAGTTCTACGGGCTCTCGGAGAAATGACGA 3589
 QY 4241 GTGCCCAAGGACCTGAATGACCTGTGCCTTATTTGAATTAACCAATCAGTTGCTTTC 4300
 Db 3590 GTGGACCCAGGATAGGGCCAAACCCGTCAACCCAGATCGTCAGGCCGAGGCTTGGGGTAG 3649
 QY 4301 TCGCTTCTGTTCCGCGCTTCTGCTCCCGAGCTCAATAAAGAGAGCCCAACCCCTCAC 4360
 Db 3650 AGCAGACTGTGGCTTCACTCCGAGTCTTACCAGCAAGGGTCTGTCTGCGCCACCATCT 3709
 QY 4361 TCGCGCGCCAGTCTCCAGATAGACTGCTGCGCCGGGTACCGGTGTTCTCAATAAACCC 4420
 Db 3710 CTATGAGATCTTGCTAGGGAAGGCCACCTTGTATGCCGTGCTGGTCAAGTGCCTCGTCT 3769
 QY 4421 TCTTGCAAGTGCATCCGACTCGTGTGCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGAT 4480
 Db 3770 GATGGCCATGG-----TCAAGAGAAAGGATTCAGAGGCTAGTCACTGACGCCA 3821
 QY 4481 TGACTACCGCTCAGCGGGTCTTTTCAGTTTCTCCACCTACACAGCTCTCACTAACATTC 4540
 Db 3822 AGCTTATCGATAAATAAAGATTTTATTTAGTCTCCAGAAAGGGGGGAAATGAAGAC 3881
 QY 4541 CTGATGTGCGCAGGACTCCGTACGCCCGT---TTTTGTTTATTAATAAATGCAAGAA 4597
 Db 3882 CCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGAAGGCATGGAATA 3941
 QY 4598 CAGTGTTCCTTCAAGCCAGACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAG 4657
 Db 3942 CATACTGAGATAGAGAAGTTCAAGTCAAGTTAGGAACAGAGAGACAGAGATATGG 4001
 QY 4658 GGCCTGTGAGCTATCTGCCACACGACTTTTAAAGATTTTATGCTCTCGGATGAGGA 4717
 Db 4002 GCCAAACAGGATATCTGTGTAAAGATTTCTGCCCCGGCTCAGGGCCAAAGACAGATGG 4061
 QY 4718 TTTAGTCAATCTATCTCTCTATTTTGTGCTTCTCGTATTTTAAATTTCTAGTTTG 4777
 Db 4062 TCCCAGATGCGGTCCCGCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTG 4121
 QY 4778 CACTCCCTTCTGAGAGACGGGATGCGAGATAGTTAATACTCTGAGGGCAGGCTCT 4837
 Db 4122 CCCCAGGACCTGAATAATGACCTGTGCTTATTTG-AACTAACCAATCAGTTGCGTCT 4180
 QY 4838 GTGAAAGGTGCTGCGGCTCAGTGTGAGATTTGGCATATAAAGGGGTCTGCCCCCTGT 4897
 Db 4181 CGCTTCTGTTCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACT 4240
 QY 4898 GTACAGACAGATCGGAATCTAGATGCACTACAGAGTCCCGCGGTCC----- 4947
 Db 4241 CGGCGCGCCAGTCTCCGATAGACTGCTGCGCCGGGTACCCGCTGTATCCAAATAAACCCCT 4300
 QY 4948 -----GGGGCTCTGACTCAGGGCATCTTGTGCTAGAGATCTCTACCGCGGACGA 4999
 Db 4301 CTTGCAAGTTGCAATCCGACTTGTGCTGCTGCTTGGAGGGGTCTCTCTGAGTGAT 4360
 QY 5000 TCGTGGCCGGGTACCGAGC-----TCGAATTCGTAATCATGCTCATAGCTGTTTCCCTGT 5053
 Db 4361 TGAATACCGCTCAGCGGGGTCTTTTCAGTATTCGTAATCATGCTCATAGCTGTTTCCCTGT 4420
 QY 5054 GTGAATTTGTTATCCGCTCACAATTCACACAACTACAGCGCGGAGAGCATAAAGTGTAA 5113
 Db 4421 GTGAATTTGTTATCCGCTCACAATTCACACAACTACAGCGCGGAGCATAAAGTGTAA 4480

QY 5114 AGCCTGGGTGCTTAATGAGTGAGCTAACTCACTAATTAATGCTTGGTTCGCTCACTGCCCCC 5173
 Db 4481 AGCCTGGGTGCTTAATGAGTGAGCTAACTCACTAATTAATGCTTGGTTCGCTCACTGCCCCC 4540
 QY 5174 TTTTCAGTTCGGGAAACCTGTGTCGAGTGTGATTAATGAATCGGCGCAACGGCGGGGAG 5233
 Db 4541 TTTTCAGTTCGGGAAACCTGTGTCGAGTGTGATTAATGAATCGGCGCAACGGCGGGGAG 4600
 QY 5234 AGGCGGTTTGGTATTTGGGCGCTTTCGCTTCTCGCTCACTGACTCGTCTGCTGCTCGGT 5293
 Db 4601 AGGCGGTTTGGCTATTTGGGCGCTTTCGCTTCTCGCTCACTGACTCGTCTGCTGCTCGGT 4660
 QY 5294 CTTTCCGCTGCGCGCAGCGGTATCAGCTCACTCAAAAGCGGTAAATACCGTTATCCACAGA 5353
 Db 4661 CTTTCCGCTGCGCGCAGCGGTATCAGCTCACTCAAAAGCGGTAAATACCGTTATCCACAGA 4720
 QY 5354 ATCAGGGGATTAACGAGGAAAGAAATGTGAGCAAAAGCGCAAAAGCGCAGGAACCG 5413
 Db 4721 ATCAGGGGATTAACGAGGAAAGAAATGTGAGCAAAAGCGCAGGAACCG 4780
 QY 5414 TAAAAAGCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCTCGTACGAGCATCACAA 5473
 Db 4781 TAAAAAGCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCTCGTACGAGCATCACAA 4840
 QY 5474 AAATCGACGCTCAAGTCAGAGTGCAGAAACCCGACAGGACTATAAAGATACAGGCGTT 5533
 Db 4841 AAATCGACGCTCAAGTCAGAGTGCAGAAACCCGACAGGACTATAAAGATACAGGCGTT 4900
 QY 5534 TCCCTCTGGAAGCTCCCTGTCGCTTCCGACCTTCCGACCTTCCGCTTACCGGATACCT 5593
 Db 4901 TCCCTCTGGAAGCTCCCTGTCGCTTCCGACCTTCCGACCTTACCGGATACCT 4960
 QY 5594 GTCGCCCTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACTGCTGAGTATCT 5653
 Db 4961 GTCGCCCTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCACTGCTGAGTATCT 5020
 QY 5654 CAGTTCCGCTGAGTGTGCTCCAAAGCTGGGCTGTGTCACGAAACCCCGCTTCAGCC 5713
 Db 5021 CAGTTCCGCTGAGTGTGCTCCAAAGCTGGGCTGTGTCACGAAACCCCGCTTCAGCC 5080
 QY 5714 CGACCGCTGCGCTTATCCGTTAACTATCGTCTGAGTCCAAACCCCGTTAAGACAGACTT 5773
 Db 5081 CGACCGCTGCGCTTATCCGTTAACTATCGTCTGAGTCCAAACCCCGTTAAGACAGACTT 5140
 QY 5774 ATCGCCACTGCGCAGCAGCCACTGGTAAAGGATTTAGCAGAGCGGATGTAGGGCGGTG 5833
 Db 5141 ATCGCCACTGCGCAGCAGCCACTGGTAAAGGATTTAGCAGAGCGGATGTAGGGCGGTG 5200
 QY 5834 TACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAGACAGATATTTGGTAT 5893
 Db 5201 TACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAAGACAGATATTTGGTAT 5260
 QY 5894 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGTGGTAGCTCTTGTATCCGGCAA 5953
 Db 5261 CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAAGTGGTAGCTCTTGTATCCGGCAA 5320
 QY 5954 ACAAAACCCCGCTGCTAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6013
 Db 5321 ACAAAACCCCGCTGCTAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5380
 QY 6014 AAAAGGATCTCAAGAAAGATCTTTCATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGA 6073
 Db 5381 AAAAGGATCTCAAGAAAGATCTTTCATCTTTTCTACGGGTCTGACGCTCAGTGGAAACGA 5440
 QY 6074 AAATCAGCTTAAGGATTTTGGTTCATGAGATTTCAAAAAGGATCTTCACTAGATCTCT 6133
 Db 5441 AAATCAGCTTAAGGATTTTGGTTCATGAGATTTCAAAAAGGATCTTCACTAGATCTCT 5500
 QY 6134 TTTAAATTAATAATCAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGA 6193
 Db 5501 TTTAAATTAATAATCAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGGTCTGA 5560
 QY 6194 CAGTTACCAATGCTTTAATCAGTGGGACCTATCTCAGCGATCTGTCTATTTTCGTTTCATC 6253

Db 5561 CAGTTACCAATGCTTAATCAGTGGACACCTATCTCAGCGATCTGTCTATTTGGTTTCATC 5620
Qy 6254 CATAGTTCCTGACTCCCGTCTGTAGATAAATACAGTACGGAGGGCTTACCATCTGG 6313
Db 5621 CATAGTTCCTGACTCCCGTCTGTAGATAAATACAGTACGGAGGGCTTACCATCTGG 5680
Qy 6314 CCCAGTGTCTCAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTTATCAGCAAT 6373
Db 5681 CCCAGTGTCTCAATGATACCGGAGACCCAGCTCACCAGCTCCAGATTTATCAGCAAT 5740
Qy 6374 AAACAGCCAGCCGGAAGGGCCGAGCGAGAGTGGCTCTCAACTTTTATCCGCTCCAT 6433
Db 5741 AAACAGCCAGCCGGAAGGGCCGAGCGAGAGTGGCTCTCAACTTTTATCCGCTCCAT 5800
Qy 6434 CCAGTCTATTAATTTGTCGCGGAAGCTTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTCGC 6493
Db 5801 CCAGTCTATTAATTTGTCGCGGAAGCTTAGAGTAAGTAGTTTCGCGAGTTAATAGTTTCGC 5860
Qy 6494 CAAGTTGTTGCCATTTGCTACAGGC - TCGTGGTGTACGCTCGCTGCTGTTGGTATGGCTTC 6552
Db 5861 CAAGTTGTTGCCATTTGCTACAGGCATCGTGGTGTACGCTCGCTGCTGTTGGTATGGCTTC 5920
Qy 6553 ATTGAGTCCCGTTCCCAACCATCAAGCGAGTTACATGATCCGCCATGTTGTGCAAAA 6612
Db 5921 ATTGAGTCCCGTTCCCAACCATCAAGCGAGTTACATGATCCGCCATGTTGTGCAAAA 5980
Qy 6613 AGCGTTAGCTTCCTTCGCTCCTCCGATCGTTGTGAGAGTAAGTTGGCGCGAGTTTATC 6672
Db 5981 AGCGTTAGCTTCCTTCGCTCCTCCGATCGTTGTGAGAGTAAGTTGGCGCGAGTTTATC 6040
Qy 6673 ACTCATGTTATGACGACACTGCATAATTCCTTACTGTCTATGTCATCCGTAAGATGCTT 6732
Db 6041 ACTCATGTTATGACGACACTGCATAATTCCTTACTGTCTATGTCATCCGTAAGATGCTT 6100
Qy 6733 TTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCAACGAG 6792
Db 6101 TTCTGTGACTGGTGAAGTACTCAACCAAGTCAATCTGAGAAATAGTATGCGGCAACGAG 6160
Qy 6793 TTGCTCTTGGCGGCTCAATACCGGATAATACCGGCAATACCGGCAATACCGGCAATACCGG 6852
Db 6161 TTGCTCTTGGCGGCTCAATACCGGATAATACCGGCAATACCGGCAATACCGGCAATACCGG 6220
Qy 6853 GCTCATCATGGAACCGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTTGAG 6912
Db 6221 GCTCATCATGGAACCGTTCTTCGGGCGGAAACTCTCAAGGATCTTACCGCTGTTGAG 6280
Qy 6913 ATCCAGTTCGATGTAAACCACTCGTGCACCCAACTCTTCAAGATCTTTCATCTTCAAC 6972
Db 6281 ATCCAGTTCGATGTAAACCACTCGTGCACCCAACTCTTCAAGATCTTTCATCTTCAAC 6340
Qy 6973 CAGGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAGGGAATAAGGC 7032
Db 6341 CAGGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCGCAAAAGGGAATAAGGC 6400
Qy 7033 GACACGGAATGTTGAATCTATCTATCTCTTCTTTTCAATTAATTAATTAATTAATTAATCA 7092
Db 6401 GACACGGAATGTTGAATCTATCTATCTCTTCTTTTCAATTAATTAATTAATTAATCA 6460
Qy 7093 GGGTTATGTTCTATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAATAGG 7152
Db 6461 GGGTTATGTTCTATGAGCGGATACATATTTGAATGATTTAGAAAAATAAACAATAGG 6520
Qy 7153 GGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGCTCTAAGAAACCATTAATCAT 7212
Db 6521 GGTTCGCGCACATTTCCCGAAAAGTGCCACCTGAGCTCTAAGAAACCATTAATCAT 6580
Qy 7213 GACATTAACCTATAAAAAATAGCGGTATCAGAGGCGCTTTTCGTCGCGGTTTCGGTGA 7272
Db 6581 GACATTAACCTATAAAAAATAGCGGTATCAGAGGCGCTTTTCGTCGCGGTTTCGGTGA 6640
Qy 7273 TGACGGTGAACCTCTGACATGACGCTCCCGAGACGGTCAAGCTTGTCTGTGAAGC 7332

Db 6641 TGACGGTGAACCTCTGACATGACGCTCCCGAGACGGTCAAGCTTGTCTGTAAAGC 6700
Qy 7333 GGATCCCGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGTTGTTGGCGGTTGTCGGGG 7392
Db 6701 GGATCCCGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGTTGTTGGCGGTTGTCGGGG 6760
Qy 7393 CTGCTTAACTATTCGGGATCAGAGCAGATTTGTTACTGAGAGTGCACCATATGCGGTGTA 7452
Db 6761 CTGCTTAACTATTCGGGATCAGAGCAGATTTGTTACTGAGAGTGCACCATATGCGGTGTA 6820
Qy 7453 AATACCCACAGATGCGTAAGGAGAAAAATACCGCATCAGGGCGCATTCGCCATTCAGGCT 7512
Db 6821 AATACCCACAGATGCGTAAGGAGAAAAATACCGCATCAGGGCGCATTCGCCATTCAGGCT 6880
Qy 7513 GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGTATTTACGCCAGCTGGCGAA 7572
Db 6881 GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGTATTTACGCCAGCTGGCGAA 6940
Qy 7573 AGGGGATGTCCTCAAGGCGATTAAGTTGGGTAACCGCAGGGTTTCCAGTCAAGAG 7632
Db 6941 AGGGGATGTCCTCAAGGCGATTAAGTTGGGTAACCGCAGGGTTTCCAGTCAAGAG 7000
Qy 7633 TTGTAACACGAGCCAGTGCC 7654
Db 7001 TTGTAACACGAGCCAGTGCC 7022

RESULT 6
US-10-471-065-20
; Sequence 20, Application US/10471065
; Publication No. US20040197855A1
; GENERAL INFORMATION:
; APPLICANT: Prof. Dr. Wiesmüller, Lisa
; TITLE OF INVENTION: Test system for the determination of genotoxicities
; FILE REFERENCE: P59532
; CURRENT APPLICATION NUMBER: US/10/471,065
; CURRENT FILING DATE: 2003-09-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20
; LENGTH: 9320
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: Plasmid
; OTHER INFORMATION: p5-Puro-CMV-(N'-EGFP)-CMV-Red-(EGFP-EU)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1592)
; OTHER INFORMATION: Retroviral vector p5NM
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3374)..(3392)
; OTHER INFORMATION: Retroviral vector p5NM
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5527)..(9320)
; OTHER INFORMATION: Retroviral vector p5NM
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1617)..(2216)
; OTHER INFORMATION: Puromycin resistance gene from pRetroOn (Clontech,
; OTHER INFORMATION: Palo Alto, CA, USA)
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (2267)..(2848)
; OTHER INFORMATION: CMV promoter from pEGFP-N1 (Clontech, Palo Alto,
; OTHER INFORMATION: CA, USA)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (2906)..(3348)
; OTHER INFORMATION: N'-EGFP, derived from EGFP from pEGFP-N1
; OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)


```

; FEATURE:
; NAME/KEY: promoter
; LOCATION: (3411)..(3992)
; OTHER INFORMATION: CMV promoter from pDsRed1-N1 (Clontech, Palo Alto, CA, USA)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4038)..(4718)
; OTHER INFORMATION: Red from pDsRed1-N1 (Clontech, Palo Alto, CA, USA)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4766)..(5508)
; OTHER INFORMATION: EGFP-EJ, derived from EGFP from pEGFP-N1
; OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)
US-10-471-065-20

Query Match      43.8%; Score 3349; DB 18; Length 9320;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 3570; Conserved 0; Mismatches 150; Indels 61; Gaps 7;

QY 3772 GCGATCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 3831
DB 3772 GCGATCCGGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 3831
QY 5541 GGGGCCCGCGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 5600
DB 5541 GGGGCCCGCGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGTCCAGGCTCTAGTTT 5600
QY 3832 TGACTCAACAATATACACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGA 3891
DB 3832 TGACTCAACAATATACACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGA 3891
QY 5601 TGACTCAACAATATACACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGA 5660
DB 5601 TGACTCAACAATATACACAGCTGAAGCCTATAGAGTACGAGCCATAGATAAAATAAAGA 5660
QY 3892 TTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTGTAGTTTGGCAAGC 3951
DB 3892 TTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTGTAGTTTGGCAAGC 3951
QY 5661 TTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTGTAGTTTGGCAAGC 5720
DB 5661 TTTTATTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACTGTAGTTTGGCAAGC 5720
QY 3952 TAGCTTAAGTAAGCCATTTTGCNAGGCATGGAATATACATACTGAGATAGAGAGTT 4011
DB 3952 TAGCTTAAGTAAGCCATTTTGCNAGGCATGGAATATACATACTGAGATAGAGAGTT 4011
QY 5721 TAGCTTAAGTAAGCCATTTTGCNAGGCATGGAATATACATACTGAGATAGAGAGTT 5780
DB 5721 TAGCTTAAGTAAGCCATTTTGCNAGGCATGGAATATACATACTGAGATAGAGAGTT 5780
QY 4012 CAGATCAAGGTTAGGAACAGAGACAGCAGAGATATGGGCCAAACAGGATATCTGGTGA 4071
DB 4012 CAGATCAAGGTTAGGAACAGAGACAGCAGAGATATGGGCCAAACAGGATATCTGGTGA 4071
QY 5781 CAGATCAAGGTTAGGAACAGAGACAGCAGAGATATGGGCCAAACAGGATATCTGGTGA 5840
DB 5781 CAGATCAAGGTTAGGAACAGAGACAGCAGAGATATGGGCCAAACAGGATATCTGGTGA 5840
QY 4072 AGCAGTTCCTGCCCC - GTTCAGGGCCAAAGAAAGTGGAAACAGGAGAAATATGGGCCAAAC 4130
DB 4072 AGCAGTTCCTGCCCC - GTTCAGGGCCAAAGAAAGTGGAAACAGGAGAAATATGGGCCAAAC 4130
QY 5841 AGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAAACAGGAGAAATATGGGCCAAAC 5900
DB 5841 AGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAAACAGGAGAAATATGGGCCAAAC 5900
QY 4131 AGGATATCTGGTGAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAGAAATATGGGCCAAAG 4190
DB 4131 AGGATATCTGGTGAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAGAAATATGGGCCAAAG 4190
QY 5901 AGGATATCTGGTGAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAGAAATATGGGCCAAAG 5960
DB 5901 AGGATATCTGGTGAAGCAGTTCCTGCCCCGCTCAGGGCCAAAGAAAGTGGAGAAATATGGGCCAAAG 5960
QY 4191 ATGCGGTCCCGCCCTCAGCAGTTCCTGAGAGAACCATCAGATGTTTCCAGGGTGCCCAAG 4250
DB 4191 ATGCGGTCCCGCCCTCAGCAGTTCCTGAGAGAACCATCAGATGTTTCCAGGGTGCCCAAG 4250
QY 5961 ATGCGGTCCCGCCCTCAGCAGTTCCTGAGAGAACCATCAGATGTTTCCAGGGTGCCCAAG 6020
DB 5961 ATGCGGTCCCGCCCTCAGCAGTTCCTGAGAGAACCATCAGATGTTTCCAGGGTGCCCAAG 6020
QY 4251 GACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCGCTTCTGCTTCTGT 4310
DB 4251 GACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCGCTTCTGCTTCTGT 4310
QY 6021 GACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCGCTTCTGCTTCTGT 6080
DB 6021 GACCTGAATAGACCTGTGCTTATTTGAACATAACCAATCAGTTCGCTTCTGCTTCTGT 6080
QY 4311 TCGCGGCTTCTGCTCCCGAGTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCC 4370
DB 4311 TCGCGGCTTCTGCTCCCGAGTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCC 4370
QY 6081 TCGCGGCTTCTGCTCCCGAGTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCC 6140
DB 6081 TCGCGGCTTCTGCTCCCGAGTCAATAAAGAGCCCAACCCCTCAGTCCGGCGGCC 6140
QY 4371 AGTCTCCGATAGATCGCTGCGCGGTTACCGGTCTCTCAATAAACCCCTCTTCAGT 4430
DB 4371 AGTCTCCGATAGATCGCTGCGCGGTTACCGGTCTCTCAATAAACCCCTCTTCAGT 4430
QY 6141 AGTCTCCGATAGATCGCTGCGCGGTTACCGGTCTCTCAATAAACCCCTCTTCAGT 6200
DB 6141 AGTCTCCGATAGATCGCTGCGCGGTTACCGGTCTCTCAATAAACCCCTCTTCAGT 6200
QY 4431 GCATCCGATCGTGGTCTCGCTGTTCTCTTTGGAGGGTCT - CTCTGAGTGAATGACTACCC 4489
DB 4431 GCATCCGATCGTGGTCTCGCTGTTCTCTTTGGAGGGTCT - CTCTGAGTGAATGACTACCC 4489
QY 6201 GCATCCGATCGTGGTCTCGCTGTTCTCTTTGGAGGGTCTCTCTGAGTGAATGACTACCC 6260
DB 6201 GCATCCGATCGTGGTCTCGCTGTTCTCTTTGGAGGGTCTCTCTGAGTGAATGACTACCC 6260
QY 4490 GTACG - GGGGCTTTTCACTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 4548
DB 4490 GTACG - GGGGCTTTTCACTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 4548
QY 6261 GTACG - GGGGCTTTTCACTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 6320
DB 6261 GTACG - GGGGCTTTTCACTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 6320
QY 4549 CCGCAGGAGTCCGTTCAGCCCGGTTTCTGTTTATAATAAAATGCAAGAACAGTGTTCCT 4608
DB 4549 CCGCAGGAGTCCGTTCAGCCCGGTTTCTGTTTATAATAAAATGCAAGAACAGTGTTCCT 4608

```

```

6321 CCGCAGGAGTCCGTTCAGCCCGGTTTGTGTTTATAATAAAATGCAAGAACAGTGTTCCT 6380
QY 4609 TCAAGCCAGACTACATCTCTGACTCTCGGCTTTATAAAAGAAAGTGTGAAGGGCTCTGTGA 4668
DB 4609 TCAAGCCAGACTACATCTCTGACTCTCGGCTTTATAAAAGAAAGTGTGAAGGGCTCTGTGA 4668
QY 6381 TCAAGCCAGACTACATCTCTGACTCTCGGCTTTATAAAAGAAAGTGTGAAGGGCTCTGTGA 6440
DB 6381 TCAAGCCAGACTACATCTCTGACTCTCGGCTTTATAAAAGAAAGTGTGAAGGGCTCTGTGA 6440
QY 4669 CTATCTGCCACACAGACTTTTAAAGATTTTATGCTCTCTGAGATGAGGATTTAGTCAATC 4728
DB 4669 CTATCTGCCACACAGACTTTTAAAGATTTTATGCTCTCTGAGATGAGGATTTAGTCAATC 4728
QY 6441 CTATCTGCCACACAGACTTTTAAAGATTTTATGCTCTCTGAGATGAGGATTTAGTCAATC 6499
DB 6441 CTATCTGCCACACAGACTTTTAAAGATTTTATGCTCTCTGAGATGAGGATTTAGTCAATC 6499
QY 4729 TATCTCTGCTATTTTGTGCTCTCTCGATTTTAAATTTTCTAGTTTGCACCTCCCTCC 4788
DB 4729 TATCTCTGCTATTTTGTGCTCTCTCGATTTTAAATTTTCTAGTTTGCACCTCCCTCC 4788
QY 6500 TATCTCTGCTATTTTGTGCTCTCTCGATTTTAAATTTTCTAGTTTGCACCTCCCTCC 6559
DB 6500 TATCTCTGCTATTTTGTGCTCTCTCGATTTTAAATTTTCTAGTTTGCACCTCCCTCC 6559
QY 4789 TGAGAGCAGCGGATTTGAGAGTGTAACTCTGAGGGCAGGCTTCTGTGAAAAGGTT 4848
DB 4789 TGAGAGCAGCGGATTTGAGAGTGTAACTCTGAGGGCAGGCTTCTGTGAAAAGGTT 4848
QY 6560 TGAGAGCAGCGGATTTGAGAGTGTAACTCTGAGGGCAGGCTTCTGTGAAAAGGTT 6619
DB 6560 TGAGAGCAGCGGATTTGAGAGTGTAACTCTGAGGGCAGGCTTCTGTGAAAAGGTT 6619
QY 4849 GCTCGGGCTCAGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACACAGAC 4908
DB 4849 GCTCGGGCTCAGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACACAGAC 4908
QY 6620 GCTCGGGCTCAGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACACAGAC 6679
DB 6620 GCTCGGGCTCAGTGTGAGATTTTGCATATAAAAGGGTCTCTGCCCTGTGTACACAGAC 6679
QY 4909 TCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCAT 4968
DB 4909 TCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCAT 4968
QY 6680 TCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCAT 6739
DB 6680 TCGGAATCTAGAGTGCATCTCAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCAT 6739
QY 4969 CTTTCCCTAGAGATCTCTACCGCGGACGATCTGTGGCGGGTACGAGCTCGAATCTGT 5028
DB 4969 CTTTCCCTAGAGATCTCTACCGCGGACGATCTGTGGCGGGTACGAGCTCGAATCTGT 5028
QY 6740 CTTTCCCTAGAGATCTCTACCGCGGACGATCTGTGGCGGGTACGAGCTCGAATCTGT 6799
DB 6740 CTTTCCCTAGAGATCTCTACCGCGGACGATCTGTGGCGGGTACGAGCTCGAATCTGT 6799
QY 5029 AATCATGTCATAGTGTGTTTC - - - - - CTGTGTGAAATGTTTATCCGCTCACA 5075
DB 5029 AATCATGTCATAGTGTGTTTC - - - - - CTGTGTGAAATGTTTATCCGCTCACA 5075
QY 6800 GGGTGTGCTGGGCTCTATATCCCGACATCACCAGTGGGGAAGATCGGGCTCGCCACTTC 6859
DB 6800 GGGTGTGCTGGGCTCTATATCCCGACATCACCAGTGGGGAAGATCGGGCTCGCCACTTC 6859
QY 5076 ATTCCACACAAACATACGAGCCGAGAGCATAAAGTGTAAAGCCTGGGGTGCCTTAAGT 5134
DB 5076 ATTCCACACAAACATACGAGCCGAGAGCATAAAGTGTAAAGCCTGGGGTGCCTTAAGT 5134
QY 6860 GGGCTCATGAGCGCTGTTTTCGGCGTGGTATGTTGGCAGGCCCTGCGCGGGGACTG 6919
DB 6860 GGGCTCATGAGCGCTGTTTTCGGCGTGGTATGTTGGCAGGCCCTGCGCGGGGACTG 6919
QY 5135 - - - - - GAGCTAATCACAATAT 5152
DB 5135 - - - - - GAGCTAATCACAATAT 5152
QY 6920 TTGGGCGCATCTCTCTGATGACACATCTCTTGGCGCGCGGTGCTCAACGGCTCAAC 6979
DB 6920 TTGGGCGCATCTCTCTGATGACACATCTCTTGGCGCGCGGTGCTCAACGGCTCAAC 6979
QY 5153 TCGGTTGCGCTCAGTCCCGCTTTCAGTTCGGGAAACCTGTCGTGAGCTGCTCAATTAAG 5212
DB 5153 TCGGTTGCGCTCAGTCCCGCTTTCAGTTCGGGAAACCTGTCGTGAGCTGCTCAATTAAG 5212
QY 6980 CTACTACTGGGCTGCTTCTTAATGAGAGTGCATTAAGGAGAGAGGCTCTGCAATTAAG 7039
DB 6980 CTACTACTGGGCTGCTTCTTAATGAGAGTGCATTAAGGAGAGAGGCTCTGCAATTAAG 7039
QY 5213 AATCGGCCAAACGCGCGGGAGAGGCGGTTTGGTATTTGGGCGCTCTTCCGCTTCTCGCT 5272
DB 5213 AATCGGCCAAACGCGCGGGAGAGGCGGTTTGGTATTTGGGCGCTCTTCCGCTTCTCGCT 5272
QY 7040 AATCGGCCAAACGCGCGGGAGAGGCGGTTTGGTATTTGGGCGCTCTTCCGCTTCTCGCT 7099
DB 7040 AATCGGCCAAACGCGCGGGAGAGGCGGTTTGGTATTTGGGCGCTCTTCCGCTTCTCGCT 7099
QY 5273 CACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5332
DB 5273 CACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5332
QY 7100 CACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7159
DB 7100 CACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7159
QY 5333 GGTATATCGGTTATCCACAGATCAGGGATACCCAGAGAAACATGTGAGCAAAAGG 5392
DB 5333 GGTATATCGGTTATCCACAGATCAGGGATACCCAGAGAAACATGTGAGCAAAAGG 5392
QY 7160 GGTATATCGGTTATCCACAGATCAGGGATACCCAGAGAAACATGTGAGCAAAAGG 7219
DB 7160 GGTATATCGGTTATCCACAGATCAGGGATACCCAGAGAAACATGTGAGCAAAAGG 7219
QY 5393 CCAGCAAAAGGCGCAGGAAACCGTAAAGGCGGTTTGGCGGTTTTCATAGGCTCCG 5452
DB 5393 CCAGCAAAAGGCGCAGGAAACCGTAAAGGCGGTTTGGCGGTTTTCATAGGCTCCG 5452
QY 7220 CCAGCAAAAGGCGCAGGAAACCGTAAAGGCGGTTTGGCGGTTTTCATAGGCTCCG 7279
DB 7220 CCAGCAAAAGGCGCAGGAAACCGTAAAGGCGGTTTGGCGGTTTTCATAGGCTCCG 7279
QY 5453 CCCCCCTGACGAGATCAAAAATTCAGCTCAAGTTCAGAGTGGCGGAAACCCGACAGG 5512
DB 5453 CCCCCCTGACGAGATCAAAAATTCAGCTCAAGTTCAGAGTGGCGGAAACCCGACAGG 5512
QY 7280 CCCCCCTGACGAGATCAAAAATTCAGCTCAAGTTCAGAGTGGCGGAAACCCGACAGG 7339
DB 7280 CCCCCCTGACGAGATCAAAAATTCAGCTCAAGTTCAGAGTGGCGGAAACCCGACAGG 7339
QY 5513 ACTATTAAGATACAGGGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGAC 5572
DB 5513 ACTATTAAGATACAGGGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGAC 5572
QY 7340 ACTATTAAGATACAGGGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGAC 7399
DB 7340 ACTATTAAGATACAGGGGTTTCCCTGGAAGCTCCCTCGTGGCTCTCTGTTTCCGAC 7399
QY 5573 CCGTGGCTTACCGGATACCTGCTGCGCTTCTCCCTTTCGGGAAGCGTGGCGCTTCTCA 5632
DB 5573 CCGTGGCTTACCGGATACCTGCTGCGCTTCTCCCTTTCGGGAAGCGTGGCGCTTCTCA 5632

```

Db 7400 CCTGCCGCTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCA 7459
Qy 5633 TAGCTACGCTGTAGGTATCTCAGTTCCGTTGAGTCCGTTCCCTCCAAGCTGGGCTGTGT 5692
Db 7460 ATGCTCACGCTGTAGGTATCTCAGTTCCGTTGAGTCCGTTCCCTCCAAGCTGGGCTGTGT 7519
Qy 5693 GCAGAAACCCCGCTTCAGCCCGGACCGCTCGGCTTATCCGGTAACTATCGTCTTGAGTC 5752
Db 7520 GCAGAAACCCCGCTTCAGCCCGGACCGCTCGGCTTATCCGGTAACTATCGTCTTGAGTC 7579
Qy 5753 CAACCCGGTAAGACACAGCTTATCGCCACTCGCAGCAGCCACTGGTAAACAGATTAGCAG 5812
Db 7580 CAACCCGGTAAGACACAGCTTATCGCCACTCGCAGCAGCCACTGGTAAACAGATTAGCAG 7639
Qy 5813 AGCAGGATAGTAGCCGCTGTCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGCTACAC 5872
Db 7640 AGCAGGATAGTAGCCGCTGTCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGCTACAC 7699
Qy 5873 TAGAAGNACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAT 5932
Db 7700 TAGAAGNACAGTATTTGGTATCTCGGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAT 7759
Qy 5933 TGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTGCA 5992
Db 7760 TGGTAGCTCTTGATCCGGCAAAACAAACCCAGCTGGTAGCGGTGTTTTTTTGTGCA 7819
Qy 5993 GCAGCAGATTACGGCGCAGAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTCTACGGG 6052
Db 7820 GCAGCAGATTACGGCGCAGAAAAAGGATCTCAAGAAAGATCCTTTGATCTTTCTACGGG 7879
Qy 6053 GTCTGAGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAA 6112
Db 7880 GTCTGAGCTCAGTGGAAACGAAAACTCAGTTAAGGATTTTGTGTCATGAGATTATCAAA 7939
Qy 6113 AAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAATTTTAAATCAATCTAAGTAT 6172
Db 7940 AAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAATTTTAAATCAATCTAAGTAT 7999
Qy 6173 ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCATCTCTCAGC 6232
Db 8000 ATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGACCATCTCTCAGC 8059
Qy 6233 GATCTGTCTATTTCGTTTCAATCAGTTGCTGATCCTCCCGTCTGTGAGTAACTACGAT 6292
Db 8060 GATCTGTCTATTTCGTTTCAATCAGTTGCTGATCCTCCCGTCTGTGAGTAACTACGAT 8119
Qy 6293 ACGGAGGGCTTACCATCTGCCCCAGTCTGCAATGATACCGCAGACCCACGCTCAC 6352
Db 8120 ACGGAGGGCTTACCATCTGCCCCAGTCTGCAATGATACCGCAGACCCACGCTCAC 8179
Qy 6353 GGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCAGAGAGTGTCTCC 6412
Db 8180 GGCTCCAGATTATCAGCAATAAACCCAGCCAGCCGGAAGGCGCAGAGAGTGTCTCC 8239
Qy 6413 TGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTAG 6472
Db 8240 TGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTTCCGGGAAGCTAGAGTAAGTAG 8299
Qy 6473 TTCCGCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTGTACAGCC-TGTTGTTGTACAG 6531
Db 8300 TTCCGCGAGTTAATAGTTTGGCAACGTTGTGCAATGCTGTACAGCCATCGTGGTGTACAG 8359
Qy 6532 CTGCTGTTGGTATGCTTCAATTCAGTCCGTTCCCAACGATCAAGGGGAGTTACATG 6591
Db 8360 CTGCTGTTGGTATGCTTCAATTCAGTCCGTTCCCAACGATCAAGGGGAGTTACATG 8419
Qy 6592 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCCCTCGGCTCGTTGTGCAAG 6651
Db 8420 ATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCCCTCGGCTCGTTGTGCAAG 8479
Qy 6652 TAAGTTGGCGCAGTGTATCAGTCATGGTTATGCGAGCACTGCAATATCTTCTTACTGT 6711
Db 8480 TAAGTTGGCGCAGTGTATCAGTCATGGTTATGCGAGCACTGCAATATCTTCTTACTGT 8539

Qy 6712 CATGCCATCCGTAAAGATGCTTTTCTGTCACTGGTGAAGTACTCAACCAAGTCAATTCTGAGA 6771
Db 8540 CATGCCATCCGTAAAGATGCTTTTCTGTGAAGTACTCAACCAAGTCAATTCTGAGA 8599
Qy 6772 ATAGTGTATCGGCGGACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATATACCGGCC 6831
Db 8600 ATAGTGTATCGGCGGACCGAGTTGCTCTTGGCCCGGCTCAATACGGGATATACCGGCC 8659
Qy 6832 ACATAGCAGAACTTTAAAGTGTCTCATATTTGAAAAAGTTTCTTCGGGGCGAAAACTCTC 6891
Db 8660 ACATAGCAGAACTTTAAAGTGTCTCATATTTGAAAAAGTTTCTTCGGGGCGAAAACTCTC 8719
Qy 6892 AAGGATCTTACCGCTGTGTGAGATCCAGTTCCAGTTAAGTAAACCCACTCGTGACCAACTGTATC 6951
Db 8720 AAGGATCTTACCGCTGTGTGAGATCCAGTTCCAGTTAAGTAAACCCACTCGTGACCAACTGTATC 8779
Qy 6952 TTACAGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGC 7011
Db 8780 TTACAGATCTTTTACTTTTACCCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGC 8839
Qy 7012 CGCAAAAAAGGGAATAAGGCGCACACGAAATTTGAAATCTCATCTCTCTCTCTCTCTCTCTCA 7071
Db 8840 CGCAAAAAAGGGAATAAGGCGCACACGAAATTTGAAATCTCATCTCTCTCTCTCTCTCTCTCA 8899
Qy 7072 ATATTATTGAAGCATTTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATTTGAATGTAT 7131
Db 8900 ATATTATTGAAGCATTTTATCAGGGTTATTTGTTCTCATGAGCGGATACATATTTGAATGTAT 8959
Qy 7132 TTAGAAAAATAAACAAATAGGGTTTCCGCGCACATTTTCCCGGAAAAAGTGCCACCTGACGT 7191
Db 8960 TTAGAAAAATAAACAAATAGGGTTTCCGCGCACATTTTCCCGGAAAAAGTGCCACCTGACGT 9019
Qy 7192 CTAGAAACCATTTATTTATCATGACATTAACCTATAAAAAATAGGGTATCACGAGGCCCTT 7251
Db 9020 CTAGAAACCATTTATTTATCATGACATTAACCTATAAAAAATAGGGTATCACGAGGCCCTT 9079
Qy 7252 TCCTCTCGCGCTTTCGGTGATGACGGTGAAAACTCTGACACATGACAGTCTCCCGGAGAC 7311
Db 9080 TCCTCTCGCGCTTTCGGTGATGACGGTGAAAACTCTGACACATGACAGTCTCCCGGAGAC 9139
Qy 7312 GGTACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAAGGGCGGTGACG 7371
Db 9140 GGTACAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCCGTCAAGGGCGGTGACG 9199
Qy 7372 GGGTGTGGCGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG 7431
Db 9200 GGGTGTGGCGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG 9259
Qy 7432 AGTGACCATATGCGGTGTGAAAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAG 7491
Db 9260 AGTGACCATATGCGGTGTGAAAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAG 9319
Qy 7492 G 7492
Db 9320 G 9320

RESULT 7

US-09-982-223A-1
; Sequence 1, Application US/09982223A
; Publication No. US20030175972A1
; GENERAL INFORMATION:
; APPLICANT: Daley, George Q.
; APPLICANT: Koh, Eugene Y.
; TITLE OF INVENTION: EXPRESSION VECTORS AND USES THEREOF
; FILE REFERENCES: 13086-002001
; CURRENT APPLICATION NUMBER: US/09/982,223A
; CURRENT FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/241,879
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Fast-Seq for Windows Version 4.0

```

; SEQ ID NO 1
; LENGTH: 5782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated nucleic acid
US-09-982-223A-1

Query Match 42.0%; Score 3213; DB 10; Length 5782;
Best Local Similarity 73.3%; Pred. No. 0;
Matches 4860; Conservative 0; Mismatches 905; Indels 864; Gaps 24;

QY 552 TGAAGACCCCACTGTAGGTTTGGCAAGCTTAGTAAGTAACCCCAATTTTGCAAGCAT 611
Db 3 TGAAGACCCCACTGTAGGTTTGGCAAGCTTAGTAAGTAACCCCAATTTTGCAAGCAT 62

QY 612 GG-AAAATACATACTGAGATACAGAGATTTCAGATCAAGGTATAGGACAGA-CAGACAG 669
Db 63 GGAATAATACATACTGAGATACAGAGATTTCAGATCAAGGTATAGGACAGA-TGGAACAG 122

QY 670 CAGAAATATGGGCCAAAACAGGATATCTGTGGTAAGCAGTTCTCTGCCCC-GTTCAGGGCCAA 728
Db 123 CTGAATATGGGCCAAAACGGGATATCTGTGGTAAGCAGTTCTCTGCCCGGCTCAGGGCCAA 182

QY 729 GAA CAGTTTGGAA CAGGAGAAATATGGGCCAAA CAGGATATCTGTGGTAAGCAGTTCTCTGCC 788
Db 183 GAA CAGATGGAA CAGCTGAATATGGGCCAAA CAGGATATCTGTGGTAAGCAGTTCTCTGCC 242

QY 789 CCGGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCCGTCCCGCCCTCAGCAGTTTCTAG 848
Db 243 CCGGCTCAGGGCCAAAGAACAGATGGTCCCAGATGCCGTCCCGCCCTCAGCAGTTTCTAG 302

QY 849 AGAAACCAATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGCCTTATTTC 908
Db 303 AGAACCAATCAGATGTTTCCAGGGTGCCCAAGGACCTGAAATGACCCCTGTGCCTTATTTC 362

QY 909 AACTAAACCAATCAGTTCCGTTCTTCGCTTCTGTGTCGCGCGCTTCTGTCTCCCGAGCTCAAT 968
Db 363 AACTAAACCAATCAGTTCCGTTCTTCGCTTCTGTGTCGCGCGCTTCTGTCTCCCGAGCTCAAT 422

QY 969 AAAAGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTCTCGATAGACTGGGTGCGCCGG 1028
Db 423 AAAAGAGCCCAACACCCCTCACTCGGCGCGCCAGTCTCTCGATAGACTGGGTGCGCCGG 482

QY 1029 TACCGGTATTCCCAATAAAGCCTCTGCTGTTTGCATCCGAATCGTGGACTCGCTGATCC 1088
Db 483 TACCGGTATTCCCAATAAAGCCTCTGCGAG-TTGATCCGACTGTGGTCTCGCTGTTC 541

QY 1089 TTGGGAGGTCCTCTCAGATTGATGACTGCC-ACCTCGGGGTCTTTTCAATTGGAGGT 1147
Db 542 TTGGGAGGTCCTCTCAGATTGATGACTACCCGTACGCGGGGTCTTTTCAATTGGGGGC 601

QY 1148 TCCACCGAGATTGGAGACCCCTGCCACAGGACCAACCGACCCCGCCCGGAGGTAAAGC 1207
Db 602 TCGTCCGGGATTCGGGAGACCCCTGCCACAGGACCAACCGA-CCCAACCAACCGGAGTAAAGC 660

QY 1208 TGGCCAGCACTTATCTGTGTCCTCGATTGTCTAGTGCTAGACTGATGATTTATGCGC 1267
Db 661 TGGCCAGCACTTATCTGTGTCCTCGATTGTCTAGTGCTAGACTGATTTATGCGC 720

QY 1268 CTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGGGGACCCCGTGGTGAACCTGAC 1327
Db 721 CTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGGGACCCCGTGGTGAACCTGAC 780

QY 1328 GAGTTTCGGAACACCCGGCGGCAACCTTCGGAGAGCGTCCAGGGAGCTTTCGGGGCGGTTTT 1387
Db 781 GAGTTTCGGAACACCCGGCGGCAACCTTCGGAGAGCGTCCAGGGAGCTTTCGGGGCGGTTTT 836

QY 1388 TGTGCGCGGACTCAGTCTTAATAATCCGATCGTTTATAGGACTCTTTTGGTCGACCCCGCTT 1447
Db 837 TGTGCGCGGACTCAGTCTTAATAATCCGATCGTTTATAGGACTCTTTTGGTCGACCCCGCTT 896

QY 1448 AGAGGAGGATATGTGGTTCTGTGTAGGAGACGAACTAAACAGTTCCTCGGCTCCGCTC 1507

```

Db 1977 GGGACACTTCACATGAGCGAAATAATACATCGTCACTGGGACATGTTGACAGATCCATGC 2036
Qy 2587 TGGTACACAGAGAGCAGGTAAAGCTCCAAAGCTGTGATCTACTGACATCCACCCGG 2646
Db 2037 ACGTAAACTCGCAAGCCGACTGATGC-----CTTCTGAACAATGGAAGGCATTAATG 2089
Qy 2647 CACACTGTGTGCGCAAGCAGATTCAGCGGTAGCGGTAGCGGTACGACTTCACCTTCACC 2706
Db 2090 CCGTAAGCCGTGGCGGTCTGTGACCGGTGGGTGAAGACAGAAACAGCACCTCGATC--- 2146
Qy 2707 ATCAGAGCTCCAGCCAGAGGACATCGCCACCTACTACTCTGCCAGCAATATAGCCTCTAT 2766
Db 2147 -TGAGCCGCGATATTGCCAGCGTTTCAACGCGCTGTATGCGGAGATCGATCCCGTC--- 2202
Qy 2767 CGGTGTTCCGCCCAAGGACCAAGGTGGAATCAAAACGAGGTGCTCAGGATCGGGTGA 2826
Db 2203 -----GTTTTCACACGTCGTGACTGGGAAACCCCTGGCGTTA 2239
Qy 2827 TCCGGCTCTGGTGTGAGTCCGAGTCCAGGTCCAACTGGTGGAGAGCGGTGGAGGTGTTGTG 2886
Db 2240 CCAACTTAATGCGCTTGGAGGACATCCCTCTTTCGCCAGCTGGCGTAAATAGCGAAAGG 2299
Qy 2887 CAACCTGCCCGTCCCTCGCTCTGCTCCGCTCCGCTCTGCTCTGCTTCCACCATAT 2946
Db 2300 CCGCACCGATCGCCCTTCCCAAAGTTGCGCGCTGAATGGCGATGGCGTTTGCC 2359
Qy 2947 TGAATGAGTTGGGTGAGACAGCCACTGGAAAAGTCTTGAGTGATTTGGAGAAATCAT 3006
Db 2360 TGGTTTCCGGCACAGAAAGCGGTGCGGAAAGCTGGCTGGAGTGCAT----- 2407
Qy 3007 CCAGTAGCAGTAACTAATGCGCGCTCTAAGAGTATGATTTACAAATATCGGA 3066
Db 2408 --CTTCTGAGCGCGGATACTGCTGCTCCCTCAAACTGGCAGATGCGAGTTACGATG 2465
Qy 3067 GACAAAGCAAGAAACACATGTTCTGCAAAATGACAGCCTGAGACCGAAGACACCGGG 3126
Db 2466 CGCCATCTACCAACGTGACCTATCCATTAAGTCAATCCGCGTTGTTCCACGG 2525
Qy 3127 GTCTATTTTGTCAAGCCTTTACTTTCGCTTCCCTGCTTGTGTTTGTGTTGTTGGGCAAGGG 3186
Db 2526 AGAATCCGACGGTGTGTTACTCGCTCACATTTTAAATGTTGATGAAGCTGCTACAGAA 2585
Qy 3187 ACCCGGTACCGTCTCAGTGCTTAAGCCACAGAGCCAGCGCCGCGACACCAACA 3246
Db 2586 GGCAGAGCGGAATTAATTTTGTGAGCGGTTAACTCGCGCTTTCATCTG----- 2633
Qy 3247 CCGCGCCCAACCATCGCGTCCAGCCCTGCTGCTGCGCCAGAGGGCGCTCGCCAGCG 3306
Db 2634 -----TGGTGAACGGCGCTGGGTGCGTTACGGGCAAGACAGTCG 2674
Qy 3307 GCGGGGGCGCAGTGCCACACAGAGGGGCTGGACTTTCGCTTGGATCCCAACTCTGCTAC 3366
Db 2675 TTTGGGTCTTAATTTGAGCTCGAGCGCATATCTACGCGCCGGA-----GNAACCGCCTC 2730
Qy 3367 CTGCTGAATGAATCCTCTTCACTATGTTGTCTATCTCACTGCTGTTGTTCTGAGAGTG 3426
Db 2731 GCGGTGATGTTGCTGCTGAGTGAACGGAGTTATCT----- 2768
Qy 3427 AAGTTCAGAGGCGCAGAGCCCCCGCTGCTGCTGCGCCAGAGGGCGCAGAACCTCTATAAC 3486
Db 2769 -----TGAAGATCAAGATATGTCGGGATGAGCGGATTCGAGCGAAAC 2814
Qy 3487 GAGCTCAATCTAGGACGAAGAGAGGTACGATGTTTGGACAAGAGACGTGGCGCGGAC 3546
Db 2815 GGTCTGCGCTCGGGAACCGCGAATTGAATTTATGGCCACACAGAGTGGCGGCGGAC 2874
Qy 3547 CCTGAGATGGGGGAAAGCCGAGAGGAAGAACCTCTCAGGAAGCCCTGTACAATGAACCTG 3606
Db 2875 TTCCAGTT-----CAACATCAGCGCTACAGTCAACAGCAACTGATGGAACACCGCCAT 2928
Qy 3607 CAGAAGATAGATGGCGGAGCCCTACAGTCAGATGGGATGGAAGGCGAGCGCGGAGG 3666
Db 2929 CGCCATCTGCTGCAAGCGGAAAGAACCCGACATGGCTGTTTATCGACGGTTTCCATATGGGG 2988

Qy 3667 GGCAGGGGCGACGATGCGCTTTACAGGGTCTCAGTACAGCCCAAGGACACCTACGAC 3726
Db 2989 ATTGGTGGCAGCAGTCTCTGAGCCCGTCAGTATCGCGGAATTCAGCTGAGCGCCGGT 3048
Qy 3727 GCCCTTCACATGCAAGGCCCTGCCCCCTCGCTTAACCTCGACGGCGCGCGGATCCGGATTAG 3786
Db 3049 CGCTACCATTAACAGTTGGTCTGGTGTCAAAAATAATAATAAACCGGCGACGCATGCTG 3108
Qy 3787 TCCAAATTTGTTAAAGACAGGATATCAGTGTGCCAGGCTCTAGTTTGTGATCAACAATATC 3846
Db 3109 CCGTATTTCGCGTAAGGAAATCCATTATGT-----ACTATTTA 3147
Qy 3847 ACCAGCTGAAGCTTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTTAGTCTCC 3906
Db 3148 AACTCGAGCGCGCCCGCAGCACAGTGGTTCGACGATAAAATAAAGATTTTATTTAGTCTCC 3207
Qy 3907 AGAAAAAGGGGGAATCAAGAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCC 3966
Db 3208 AGAAAAAGGGGGAATCAAGAGACCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCC 3267
Qy 3967 CATTTTCAAGGCTAGG-AAAAATACATACTAGAGATAGAGAGTTTCAGATCAAGGTTAG 4025
Db 3268 CATTTTGAAGGCTAGGAAAAATACATACTAGAGATAGAGAGTTTCAGATCAAGGTCAG 3327
Qy 4026 GAAACAGA-GAGACAGAGATATGGGCCAAACAGGATATCTGTGTAGAGAGTTCTCTGCC 4084
Db 3328 GAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTAGAGAGTTCTCTGCC 3387
Qy 4085 CC-GCTCAGGGCCAAAGAACAGTTGGAAACAGAGAGATATGGGCCAAACAGGATATCTGTGG 4143
Db 3388 CCGCTCAGGGCCAAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGG 3447
Qy 4144 TAAGCAGTTCTGTCGCCCGGCTCAGGGCCAAAGACAGATGTTGCCAGATCGGTCGCCGCC 4203
Db 3448 TAAGCAGTTCTGTCGCCCGGCTCAGGGCCAAAGACAGATGTTGCCAGATCGGTCGCCGCC 3507
Qy 4204 CTACAGCATTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAAAGCACTGAAATGAC 4263
Db 3508 CTACAGCATTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCCAAAGCACTGAAATGAC 3567
Qy 4264 CTTGTGCTTATTTGAATTAACCAATCAGTTCGCTTCTCGTTCGCTTCTGTTCCGCGCTTCTG 4323
Db 3568 CTTGTGCTTATTTGAATTAACCAATCAGTTCGCTTCTCGTTCGCTTCTGTTCCGCGCTTCTG 3627
Qy 4324 CTCGCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCTCCGATAG 4383
Db 3628 CTCGCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGCCAGTCTCTCCGATAG 3687
Qy 4384 ACTGCTGCGCCCGGTTACCGGTGTTCTCAATAAACCCCTTTTGCAAGTTGCACTCCGACTCGT 4443
Db 3688 ACTGAGTCCGCGGGTACCGGTGTTATCAATAAACCCCTTTTGCAAGTTGCACTCCGACTTGT 3747
Qy 4444 GGTCTCCTGTTCTCTGGAGGGTCT-CTCTGAGTGATTTGATACCCGTCAGCGGGGCT 4502
Db 3748 GGTCTCCTGTTCTCTGGAGGGTCTCTCTGAGTGATTTGATACCCGTCAGCGGGG--- 3804
Qy 4503 TTCAGTTTCTCCCACTACACAGGTCTCACTAAATTCCTGATGTGCGCGAGGACTCCG 4562
Db 3805 ----- 3804
Qy 4563 TCAGCCCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTTCAAGCCAGACTAC 4622
Db 3805 ----- 3804
Qy 4623 ATCTGACTCTCGGCTTTTAAAGAAATGTTGAAGGGCTCTGTGAGCTATCTGCCACAG 4682
Db 3805 ----- 3804
Qy 4683 ACTTTTAAAGATTTTATGCTCTCGGATGAGGATTTAGTCAATCTATCTCTGCTATT 4742
Db 3805 ----- 3804

QY 4743 TTGCTGGCTTCCGCTATTTTAAATTTCTAGTTTGCATCCCTTCTGTGAGACGCGCA 4802
 Db 3805 ----- 3804
 QY 4803 TTGCAGAGTAGTTAATACTCTGAGGCGAGCTTCTGTGAAAGGTTGCCCTGGCTCAGTG 4862
 Db 3805 ----- 3804
 QY 4863 TGAGATTTTGGCCATAAAAGGGTCTCTGCCCTGTGTACAGACAGATCGGAATCTAGACT 4922
 Db 3805 ----- 3804
 QY 4923 GCATACTCAGAGTCCCGCGGTTTCGGGGCTCTGATCTCAGGGCATCTTTTGGCTAGAGAT 4982
 Db 3805 ----- 3804
 QY 4983 CCTCTACGGCGGACGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGTGTCATAG 5042
 Db 3805 ----- 3804
 QY 5043 CTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCACACACATACAGCGCGGAGC 5102
 Db 3805 ----- 3804
 QY 5103 ATAAAGTGTAAAGCCTGGGGTGTCTAATGAGTGAGCTAACTCACATTAATTCGGTTGCGC 5162
 Db 3805 ----- 3804
 QY 5163 TCAGTCCCGCTTTCAGTTCGGGAACCTGTGTGCGCAGCTGATTAATGAATTCGGCCAA 5222
 Db 3805 -----GTCTTTCATTCGCAATTAATGAATTCGGCCAA 3835
 QY 5223 CGCGGGGAGAGCGGTTTGGTATTTGGCGCTTTCGCTTCCGCTCACTGACTCG 5282
 Db 3836 CGCGGGGAGAGCGGTTTGGTATTTGGCGCTTTCGCTTCCGCTCACTGACTCG 3895
 QY 5283 CTGCGCTCGGCTTTCGCTGGGGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG 5342
 Db 3896 CTGCGCTCGGCTTTCGCTGGGGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGG 3955
 QY 5343 TTATCCACAGATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGGCCAGCAAAAG 5402
 Db 3956 TTATCCACAGATCAGGGGATTAACGAGGAAAGACATGTGAGCAAAAGGCCAGCAAAAG 4015
 QY 5403 GCCAGAACCGTAAAGGCGCGTTCTGGCGTTTTCATAGCTCCGCCCTGAC 5462
 Db 4016 GCCAGAACCGTAAAGGCGCGTTCTGGCGTTTTCATAGCTCCGCCCTGAC 4075
 QY 5463 GAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGA 5522
 Db 4076 GAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGA 4135
 QY 5523 TACAGGCGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTTCCGCTT 5582
 Db 4136 TACAGGCGTTTCCCGCTGGAAGTCCCTCGTGGCTCTCTGTTCCGACCTTCCGCTT 4195
 QY 5583 ACGGATACCTGTCCGCTTCTCCCTTCGGGAAGGCTGGCTTCTCATAGCTCACGC 5642
 Db 4196 ACGGATACCTGTCCGCTTCTCCCTTCGGGAAGGCTGGCTTCTCATAGCTCACGC 4255
 QY 5643 TGTAGGTATCTCAGTTCGGGTAGGTTCGTTCCGCTCAAGCTGGGCTGTGTGACGAACCC 5702
 Db 4256 TGTAGGTATCTCAGTTCGGGTAGGTTCGTTCCGCTCAAGCTGGGCTGTGTGACGAACCC 4315
 QY 5703 CCGGTTTACGCCGAGCGGCTTCCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTA 5762
 Db 4316 CCGGTTTACGCCGAGCGGCTTCCGCTTATCCGGTAACTATCGTCTTGAAGTCCAAACCGGTA 4375
 QY 5763 AGACAGCATTTATCGGCATCGCAGCAGCCTGTGTAAAGGATTTAGCAGCAGGATAT 5822
 Db 4376 AGACAGCATTTATCGGCATCGCAGCAGCCTGTGTAAAGGATTTAGCAGCAGGATAT 4435
 QY 5823 GTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTTAACCTACGGGCTACACTAGAAGACA 5882

Db 4436 GTAGGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGGCTACACTAGAAGACA 4495
 QY 5883 GTATTTGGTATCTCGCTCTGCTGAGCCAGTTACTTTCGGAAAGAGTTCGTTAGTCTCT 5942
 Db 4496 GTATTTGGTATCTCGCTCTGCTGAGCCAGTTACTTTCGGAAAGAGTTCGTTAGTCTCT 4555
 QY 5943 TGATCCGGCAAAACCAACCGCTGCTAGCGGTGGTTCGTTTGGTTCGAAAGCAGAGATT 6002
 Db 4556 TGATCCGGCAAAACCAACCGCTGCTAGCGGTGGTTCGTTTGGTTCGAAAGCAGAGATT 4615
 QY 6003 ACGCGCAGAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTTCTACCGGGTCTGACGCT 6062
 Db 4616 ACGCGCAGAAAAAGGATCTCAAGAAGATCTTTTGAATCTTTTCTACCGGGTCTGACGCT 4675
 QY 6063 CAGTGGACGAAAACTCAGGTTAAGGATTTGGTTCATGAGATTATCAAAAAGGATCTTC 6122
 Db 4676 CAGTGGACGAAAACTCAGGTTAAGGATTTGGTTCATGAGATTATCAAAAAGGATCTTC 4735
 QY 6123 ACCTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 6182
 Db 4736 ACCTAGATCTTTT-----GCGGCGCGCGCAATCAATCAATCAATCAATCAATCAATCAAT 4788
 QY 6183 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTGTCTA 6242
 Db 4789 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTTATCTCAGCGATCTGTCTA 4848
 QY 6243 TTTGCTTCTCATCATAGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6302
 Db 4849 TTTGCTTCTCATCATAGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4908
 QY 6303 TTACCATCTGGCCCGCAGTGTCAATGATACCGCAGACCCACGCTCACCAGCTCCAGAT 6362
 Db 4909 TTACCATCTGGCCCGCAGTGTCAATGATACCGCAGACCCACGCTCACCAGCTCCAGAT 4968
 QY 6363 TTATCAGCAATAAAACAGCCAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 6422
 Db 4969 TTATCAGCAATAAAACAGCCAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 5028
 QY 6423 TCCGCTCGATCAGTCTATTAATTTGCTGCGGGAAGCTAGAGTAAGTATGTCGCGAGTT 6482
 Db 5029 TCCGCTCGATCAGTCTATTAATTTGCTGCGGGAAGCTAGAGTAAGTATGTCGCGAGTT 5088
 QY 6483 AATAGTTTGGCAACGTTGTTGCCATTGCTCAGGCACTCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 6541
 Db 5089 AATAGTTTGGCAACGTTGTTGCCATTGCTCAGGCACTCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 5148
 QY 6542 GGTATGGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTTACATGATCCCCCATG 6601
 Db 5149 GGTATGGCTTCAATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTTACATGATCCCCCATG 5208
 QY 6602 TTGTCGCAAAAAGCGGTTAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 6661
 Db 5209 TTGTCGCAAAAAGCGGTTAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 5268
 QY 6662 CGAGTGTATCATCATGTTATGCGACACTGCAATTTCTTACTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6721
 Db 5269 CGAGTGTATCATCATGTTATGCGACACTGCAATTTCTTACTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5328
 QY 6722 GTAAAGTCTTTTCTGACTGGTGTGACTCAACCAAGTCAATTTCTGAGAATAGTGTATG 6781
 Db 5329 GTAAAGTCTTTTCTGACTGGTGTGACTCAACCAAGTCAATTTCTGAGAATAGTGTATG 5388
 QY 6782 CGGCGACCGAGTTGCTTTCGCGCGGTCAATACGCGGTAATACCGCGCAATAGACAGA 6841
 Db 5389 CGGCGACCGAGTTGCTTTCGCGCGGTCAACACCGGTAATACCGCGCAATAGACAGA 5448
 QY 6842 ACTTTAAAGTCTCATCATTTGGAAGAGTTCCTTCGGGCGGAAACTCTCAGGATCTTGA 6901
 Db 5449 ACTTTAAAGTCTCATCATTTGGAAGAGTTCCTTCGGGCGGAAACTCTCAGGATCTTGA 5508
 QY 6902 CGGCTGTTGAGATCCAGTTCGATTAACCCACTCGTGCACCAACTGATCTTTCAGCATCT 6961

Db 2977 GTGGTAACAGTTCTCTGCCCGGCTCAGGGCCAGAACAGATGGTCCCAAGATGGGTCC 3036
Qy CCCCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGAACCTG-AA 4258
Db CGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGAACCTGAAA 3096
Qy ATGACCCCTGTGCTGTTATTTGAACCTAACCAATCAGTTGCTCTCTGTTGCGCGGC 4318
Db ATGACCCCTGTGCTGTTATTTGAACCTAACCAATCAGTTGCTCTCTGTTGCGCGGC 3156
Qy TTCTGCTCCCGAGCTCAATAAAGAGAGCCCAACCCCTCACTCGGCGCGGCGAGTCCCTCC 4378
Db TTCTGCTCCCGAGCTCAATAAAGAGAGCCCAACCCCTCACTCGGCGCGGCGAGTCCCTCC 3216
Qy GATAGACTGCGTCCGCGGGTACCCGTTTCTCAATAAACCCTCTTGCAGTTGATCCGA 4438
Db GATAGACTGCGTCCGCGGGTACCCGTTTCTCAATAAACCCTCTTGCAGTTGATCCGA 3276
Qy CTGCTGCTCTGCTGTTCTCTTGGAGGGTCT-CTCTGAGTGATTTGATACCCGTCAGCGG 4497
Db CTGCTGCTCTGCTGTTCTCTTGGAGGGTCTCTCTGAGTGATTTGATACCCGTCAGCGG 3336
Qy GGTCTTTTCAGTTTCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGGA 4557
Db GGTCTTTTCAGTTTCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGGA 4557
Qy GGTCTTTTCAGTTTCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGGA 4557
Db GGTCTTTTCAGTTTCTCCACCTACACAGGTCTCACTAATCTCTGATGTGCCGAGGGA 4557
Qy CTCGCTCAGCCCGGTTTTTGTATTATAAATAAAGAGAGTGTCCCTTCAAGCCAG 4617
Db CTCGCTCAGCCCGGTTTTTGTATTATAAATAAAGAGAGTGTCCCTTCAAGCCAG 4617
Qy ACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGCTCTGTGAGTATCTGCC 4677
Db ACTACATCTGACTCTCGGCTTTTATAAAGAAATGTTGAAGGCTCTGTGAGTATCTGCC 4677
Qy ACACGACTTTTAAAGATTTTATGCTCTCGATGAGGGATTTAGTCAATCTATCTCTCGT 4737
Db ACACGACTTTTAAAGATTTTATGCTCTCGATGAGGGATTTAGTCAATCTATCTCTCGT 4737
Qy CTATTTTGTCTGGCTTCTCGGTTATTTAAATTTCTAGTTTGCACCTCCCTTCTCGAGAGCAC 4797
Db CTATTTTGTCTGGCTTCTCGGTTATTTAAATTTCTAGTTTGCACCTCCCTTCTCGAGAGCAC 4797
Qy GGCAGTTGACAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGTTGCTCGGCT 4857
Db GGCAGTTGACAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAGTTGCTCGGCT 4857
Qy CAGTGTGAGATTTGCCATAAAGGGGTCTGCCCTGTGTACAGACAGATCGGAATCT 4917
Db CAGTGTGAGATTTGCCATAAAGGGGTCTGCCCTGTGTACAGACAGATCGGAATCT 4917
Qy AGAGTGCACTCAGAGTCCCGGTTCCGGGGCTCTGATCTCAGGGGATCTTTTGCTTA 4977
Db AGAGTGCACTCAGAGTCCCGGTTCCGGGGCTCTGATCTCAGGGGATCTTTTGCTTA 4977
Qy GAGATCTCTACGCCGACGATCGTGGCGGGTACCGAGTCCGAATTCGTAATCATGGT 5037
Db GAGATCTCTACGCCGACGATCGTGGCGGGTACCGAGTCCGAATTCGTAATCATGGT 5037
Qy CATAGCTGTTTCTGTGTGAATTTGTTATCGCTCAACAATTCACACATACGAGCGG 5097
Db CATAGCTGTTTCTGTGTGAATTTGTTATCGCTCAACAATTCACACATACGAGCGG 5097
Qy GAACATAAAGTGTAAAGCTTGGGTTGCTTAATCAGTGAGTAACTCACATTAATTCGGT 5157
Db GAACATAAAGTGTAAAGCTTGGGTTGCTTAATCAGTGAGTAACTCACATTAATTCGGT 5157
Qy TGCCTCACTGCCCGCTTTTCAGTCCGGAACCTGTGCTGCGAGCTGCATTAATGAATCG 5217
Db TGCCTCACTGCCCGCTTTTCAGTCCGGAACCTGTGCTGCGAGCTGCATTAATGAATCG 5217
Qy GCCAACCGGCGGAGAGGGGTTTGGTATTTGGGGGCTCTTCCGCTTCTCGCTCACTG 5277
Db GCCAACCGGCGGAGAGGGGTTTGGTATTTGGGGGCTCTTCCGCTTCTCGCTCACTG 5277

Db 3544 GCCAACCGGCGGAGAGGGCGGTTTGGTATTTGGGCGCTCTTCCGCTCTCTCGCTCACTG 3603
Qy ACTCGCTCGCTCGGTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTAA 5337
Db ACTCGCTCGCTCGGTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAGAGCGGTAA 3663
Qy TACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAACATGTGAGGAAAGAGCCAGC 5397
Db TACGGTTATCCACAGATCAGGGATTAACGAGGAAAGAACATGTGAGGAAAGAGCCAGC 3723
Qy AAAAGGCGAGGAAACCGTAAAGGCGGTTGCTGGCGGTTTTTCCATAGGCTCCGCCCCC 5457
Db AAAAGGCGAGGAAACCGTAAAGGCGGTTGCTGGCGGTTTTTCCATAGGCTCCGCCCCC 3783
Qy CTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAT 5517
Db CTGACGAGCATCACAAAAATCGACGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAT 3843
Qy AAAGATACAGGCGTTCCTCCCTCGGAAAGTCTCTGCTGCTCTCTCTGTTCCGACCTGTC 5577
Db AAAGATACAGGCGTTCCTCCCTCGGAAAGTCTCTGCTGCTCTCTCTGTTCCGACCTGTC 5577
Qy AAAGATACAGGCGTTCCTCCCTCGGAAAGTCTCTGCTGCTCTCTCTGTTCCGACCTGTC 3903
Db AAAGATACAGGCGTTCCTCCCTCGGAAAGTCTCTGCTGCTCTCTCTGTTCCGACCTGTC 5577
Qy CGCTTACCGGATACCTGTCCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCT 5637
Db CGCTTACCGGATACCTGTCCGCGCTTTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCT 3963
Qy CACGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGACG 5697
Db CACGCTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGCTCAAGCTGGGCTGTGTGACG 4023
Qy AACCCCGGTTACGCCGAGCGCTGCGCTTATCCGCTTACCTGCTTCCGCTTCCGCTTCCGCT 5757
Db AACCCCGGTTACGCCGAGCGCTGCGCTTATCCGCTTACCTGCTTCCGCTTCCGCTTCCGCT 4083
Qy CGGTAAGACACGCTTATCGGCTGCTGAGGTTCTTGAAGTGGTGGCTTAACTGCTGCTGCTGCTGCT 5817
Db CGGTAAGACACGCTTATCGGCTGCTGAGGTTCTTGAAGTGGTGGCTTAACTGCTGCTGCTGCTGCT 4143
Qy GGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACTGCTGCTGCTGCTGCTGCTGCT 5877
Db GGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACTGCTGCTGCTGCTGCTGCTGCT 4203
Qy GGACAGTATTTGGTATCTGCGCTCTGCTGAGCGAGTTCCTGCGGAAAGAGTGGTA 5937
Db GGACAGTATTTGGTATCTGCGCTCTGCTGAGCGAGTTCCTGCGGAAAGAGTGGTA 4263
Qy GCTCTTATCGGCGGAAACCAACCGCTGCTGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 5997
Db GCTCTTATCGGCGGAAACCAACCGCTGCTGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4263
Qy AGATTACGCGGAGGAAAGGATCTCAAGAGATCTTTCGATCTTTCGATCTTTCGATCTTTCGATCTTTCGAT 6057
Db AGATTACGCGGAGGAAAGGATCTCAAGAGATCTTTCGATCTTTCGATCTTTCGATCTTTCGATCTTTCGAT 4383
Qy ACGTCTAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAAGAGATTTTAAAGGATTTTAAAGGAT 6117
Db ACGTCTAGTGGAAACGAAACCTCAGTTTAAAGGATTTTGGTCAAGAGATTTTAAAGGATTTTAAAGGAT 4443
Qy TCTTACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 6177
Db TCTTACCTAGATCTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 4503
Qy AGTAAACCTGCTGACAGTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 6237
Db AGTAAACCTGCTGACAGTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 4563
Qy GTCTATTTCTGTTCAATCAGTTTGGCTGACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6297
Db GTCTATTTCTGTTCAATCAGTTTGGCTGACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4623
Qy AGGGCTTACCATCTGGGCGGAGGCGGTTTGGTATTTGGGGGCTCTTCCGCTTCTCGCTCACTG 6357
Db AGGGCTTACCATCTGGGCGGAGGCGGTTTGGTATTTGGGGGCTCTTCCGCTTCTCGCTCACTG 4683

QY 6358 CAGATTATCAGCAATAAACACGACGACGCGAAGGCGGAGCGCAGAAAGTGGTCCCTGCAA 6417
 DB 4684 CAGATTATCAGCAATAAACACGACGACGCGAAGGCGGAGCGCAGAAAGTGGTCCCTGCAA 4743
 QY 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTAGAGTAAGTAGTTCG 6477
 DB 4744 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAAGCTAGAGTAAGTAGTTCG 4803
 QY 6478 CAGTTAATAGTTGCGCAAGCTTTGTCATGCTACAGC-TCGTGGTGTACGCTCGT 6536
 DB 4804 CAGTTAATAGTTGCGCAAGCTTTGTCATGCTACAGC-TCGTGGTGTACGCTCGT 4863
 QY 6537 CGTTTGGTATGGCTTCATTACGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCC 6596
 DB 4864 CGTTTGGTATGGCTTCATTACGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCC 4923
 QY 6597 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGT 6656
 DB 4924 CCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCCTCCGATCGTTGTGAGAAGTAAGT 4983
 QY 6657 TGGCCCGAGTTTATCAGTATGTTATGCGAGCACTGCATTAATTTCTTACTGTGTCATGC 6716
 DB 4984 TGGCCCGAGTTTATCAGTATGTTATGCGAGCACTGCATTAATTTCTTACTGTGTCATGC 5043
 QY 6717 CATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATTCAGATAGT 6776
 DB 5044 CATCCGTAAGATGCTTTTCTGTGACTGCTGAGTACTCAACCAAGTCAATTCAGATAGT 5103
 QY 6777 GTATGCGGACCGAGTGTCTTTCGCGCGCTCAATACGGGATATACCGGCCACATA 6836
 DB 5104 GTATGCGGACCGAGTGTCTTTCGCGCGCTCAATACGGGATATACCGGCCACATA 5163
 QY 6837 GCAGAACTTTAAAGTGTCACTATTGGAAGAGTTCCTTCGGGCGGAAACTCTCAAGA 6896
 DB 5164 GCAGAACTTTAAAGTGTCTCACTATTGGAAGAGTTCCTTCGGGCGGAAACTCTCAAGA 5223
 QY 6897 TCTTACCGTGTGAGATCCAGTTCGATGTAACCACTCGTGCACCCCACTGATCTTCAG 6956
 DB 5224 TCTTACCGTGTGAGATCCAGTTCGATGTAACCACTCGTGCACCCCACTGATCTTCAG 5283
 QY 6957 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 7016
 DB 5284 CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCGCAA 5343
 QY 7017 AAAAGGAATAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATT 7076
 DB 5344 AAAAGGAATAGGGCGACACGGAATGTTGAATCTCATCTCTTCTCTTTTCAATATT 5403
 QY 7077 ATTGAAGCATTTATCAGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTTGA 7136
 DB 5404 ATTGAAGCATTTATCAGGGTTATTGTCATGAGCGGATACATATTTGAATGATTTTGA 5463
 QY 7137 AAAATAAAACAAATAGGGGTTCCGCGACACATTTCCCGGAAAGTGCCACCTGACGCTAAG 7196
 DB 5464 AAAATAAAACAAATAGGGGTTCCGCGACACATTTCCCGGAAAGTGCCACCTGACGCTAAG 5523
 QY 7197 AAACCAATTAATCATGACATTAACCTATAAAAATAGGGGTTATCACAGCGCCCTTTTCGTC 7256
 DB 5524 AAACCAATTAATCATGACATTAACCTATAAAAATAGGGGTTATCACAGCGCCCTTTTCGTC 5583
 QY 7257 TCGCGGTTTCGGTGTATGAGCGTGAACCTCTGACACATGACGCTCCCGGAGCGGTCA 7316
 DB 5584 TCGCGGTTTCGGTGTATGAGCGTGAACCTCTGACACATGACGCTCCCGGAGCGGTCA 5643
 QY 7317 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCGCTCAGGGCGGCTCAGCGGGTG 7376
 DB 5644 CAGCTTGTCTGTAAGCGGATGCGGAGCAGACAAGCCGCTCAGGGCGGCTCAGCGGGTG 5703
 QY 7377 TTGGCGGGTGTGCGGGCTGGCTTAATCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC 7436
 DB 5704 TTGGCGGGTGTGCGGGCTGGCTTAATCTATGCGGCATCAGAGCAGATTTGACTGAGAGTGC 5763

QY 7437 ACCATATGCGGTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGCGCC 7496
 DB 5764 ACCATATGCGGTGTGAATATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGCGCC 5823
 QY 7497 ATTGCCATTTCAGGCTGCGCAACTGTTCGGAAGGCGATTCGTCGGGGCTCTTCGCTAT 7556
 DB 5824 ATTGCCATTTCAGGCTGCGCAACTGTTCGGAAGGCGATTCGTCGGGGCTCTTCGCTAT 5883
 QY 7557 TACGCCAGCTGCGGAAGGGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGCT 7616
 DB 5884 TACGCCAGCTGCGGAAGGGGATGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGCT 5943
 QY 7617 TTTCCAGTCACGACTTGTGTAACACGCGCCAGTGCC 7654
 DB 5944 TTTCCAGTCACGACTTGTGTAACACGCGCCAGTGCC 5981

RESULT 9
 US-10-789-938B-1
 ; Sequence 1, Application US/10789938B
 ; Publication No. US20050009180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yang, Lili
 ; APPLICANT: Van Pazijs, Luk
 ; APPLICANT: Baltimore, David
 ; TITLE OF INVENTION: METHOD FOR THE GENERATION OF
 ; TITLE OF INVENTION: ANTIGEN-SPECIFIC LYMPHOCYTES
 ; FILE REFERENCE: CALTE.008CP1
 ; CURRENT APPLICATION NUMBER: US/10/789,938B
 ; PRIOR APPLICATION NUMBER: 10/317,078
 ; PRIOR FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: 60/394,803
 ; PRIOR FILING DATE: 2002-07-08
 ; PRIOR APPLICATION NUMBER: 60/339,375
 ; PRIOR FILING DATE: 2001-12-10
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 6254
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: This represents a retroviral vector derived from
 ; OTHER INFORMATION: the murine stem cell virus.
 US-10-789-938B-1

Query Match 41.2%; Score 3151.6; DB 19; Length 6254;
 Best Local Similarity 71.3%; Pred. No. 0;
 Matches 5077; Conservative 0; Mismatches 889; Indels 1152; Gaps 24;

QY 552 TGAAGAGCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTCAAGGCAT 611
 DB 1 TGAAGAGCCCACTGTAGTTTGGCAAGCTAGCTTAAGTAAGCCATTTTCAAGGCAT 60
 QY 612 GGAATAATACATCACTGAGATAGAGAGTTTCAAGTTAGGAAACAGAGAGCAGCA 671
 DB 61 GGAATAATACATCACTGAGATAGAGAGTTTCAAGTTAGGAAACAGAGAGCAGC- 119
 QY 672 GAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCGCTCAGGCGCAAGAA 731
 DB 120 ----- 119
 QY 732 CAGTTGGAACAGGAGAAATATGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCG 791
 DB 120 -----AGAAATATGGGCCAAACAGGATATCTGTGTAAGCAGTTCCTGCCCG 166
 QY 792 GCTCAGGCGCAAGAACAGATGCTCCCGAGATGCGGCTCCCGCTCAGCAGTTTCTAGAGA 851
 DB 167 GCTCAGGCGCAAGAACAGATGCTCCCGAGATGCGGCTCCCGCTCAGCAGTTTCTAGAGA 226
 QY 852 ACCATCAGATGTTTCCAGGCTGCCCAAGGACCTG-AAATGACCTGTCCTTATTGAA 910

Db 227 ACATCAGATGTTTCCAGGTTGCCCAAGGACCTGAAATGACCCGTGCGCTTATTTGAA 286
Qy 911 CTAACCAATCAGTTTCGCTTCTCGTTCCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA 970
Db 287 CTAACCAATCAGTTTCGCTTCTCGTTCCTGTTTCGGCGCTTCTGCTCCCGAGCTCAATAA 346
Qy 971 AAGAGCCACAAACCCCTCAGTTCGGGGCCAGTCTCGGATAGACTGCGTCCCGGGTA 1030
Db 347 AAGAGCCACAAACCCCTCAGTTCGGGGCCAGTCTCGGATAGACTGCGTCCCGGGTA 406
Qy 1031 CCGTATTTCCCAATAAAGCCTTCTGCTTTTGCATCGAATCGTGGACTCGCTCATCTT 1090
Db 407 CCGTATTTCCCAATAAAGCCTTCTGCTTTTGCATCGAATCGTGGACTCGCTCATCTT 466
Qy 1091 GGGAGGCTCTCCTCAGATTGATTGACTGCCACCTCGGGGCTTTTCAATTTGGAGTTCC 1150
Db 467 GGGAGGCTCTCCTCAGATTGATTGACTGCCACCTCGGGGCTTTTCAATTTGGAGTTCC 526
Qy 1151 ACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAGGTAAGCTGG 1210
Db 527 ACCGAGATTGGAGACCCCTGCCAGGACCAACCGACCCCGCGGGAGGTAAGCTGG 586
Qy 1211 CCAGCACTTATCTGTCTGTC-----CGATTGCTAGTGTCTATGACTCATTTTA 1262
Db 587 CCAGGCTGCTTCTGTCTGCTCTGCTTTTGTGCGGTTTGTGCGGCACTCAATGTT 646
Qy 1263 TCGCCTCGCTCGGTACTAGTTAGTAACTAGCTCTGTATCTGGCGACCCGCTGGTGA 1322
Db 647 TCGCCTCGCTCTGTACTAGTTAGTAACTAGCTCTGTATCTGGCGACCCGCTGGTGA 706
Qy 1323 CTGACGAGTTCCGAAACACCGCGCGCAACCTTGGAGAGCTGCCAGGAGTTCCGGGGCC 1382
Db 707 CTGACGAGTTCTGAACACCGCGCGCAACCTTGGAGAGCTGCCAGGAGTTTGGGGGCC 766
Qy 1383 GTTTTGTGGCGCGACCTGAGTCTTAAATCCGATCGTTTAGGACTTTTGTGTCACCC 1442
Db 767 GTTTTGTGGCGCGACCTGAGGAAGGGAGTCAATGGAATCCGACCCCGTCA----- 819
Qy 1443 CCCTTAGAGGAGGATATGTTGTTCTGTTAGAGACAGAACTTAAACAGTTCCCGCT 1502
Db 820 -----GGATATGTTGTTCTGTTAGAGACAGAACTTAAACAGTTCCCGCT 867
Qy 1503 CCGTCTGAATTTTGTCTTTCGTTTGGACCGAAGCGCGCGCTTGT---CTGC 1559
Db 868 CCGTCTGAATTTTGTCTTTCGTTTGGAAACGAAGCGCGCTTGTCTGCTGCAGCGC 927
Qy 1560 TGCAGCATCGTTCTGTGTGTCTGTCTGATGTTTCTGTATTTGTCTGAAATATG 1619
Db 928 TGCAGCATCGTTCTGTGTGTCTGTCTGATGTTTCTGTATTTGTCTGAAATATG 987
Qy 1620 GGCCCGGCTAGACTTTTACACTCCCTTAAAGTTTGAAGTTAGGTTCACTGGAAGATGTC 1679
Db 988 GGCC-----AGACTTTTACACTCCCTTAAAGTTTGAAGTTAGGTTCACTGGAAGATGTC 1041
Qy 1680 GAGCGGATCGTCTCAACACAGTTCGTTAGATGTCAAGAAAGACGTTGGGTTTACCTTTCG 1739
Db 1042 GAGCGGATCGTCTCAACACAGTTCGTTAGATGTCAAGAAAGACGTTGGGTTTACCTTTCG 1101
Qy 1740 TCTGAGAAATGGCCAACTTTAAAGTTCGATGCGCGGAGAGCGGACCTTTAAACGAGAC 1799
Db 1102 TCTGAGAAATGGCCAACTTTAAAGTTCGATGCGCGGAGAGCGGACCTTTAAACGAGAC 1161
Qy 1800 CTCATCACCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCATGGACACCCAGACAG 1859
Db 1162 CTCATCACCAGGTTAAGATCAAGTCTTTTACCTGGCCCGCATGGACACCCAGACAG 1221
Qy 1860 GTCCCTCTACATCGTACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTTCAAGCCC 1919
Db 1222 GTCCCTCTACATCGTACCTGGGAAGCTTGGCTTTTGAACCCCTTCCCTGGGTTCAAGCCC 1281
Qy 1920 TTTGTACACCTTAAGCTTCGCTCTTCTCTCCATCCGCGCGCTCTCTCCCTTTGAA 1979
Db 1282 TTTGTACACCTTAAGCTTCGCTCTTCTCTCCATCCGCGCGCTCTCTCCCTTTGAA 1341

Qy 1980 CTTCTCTGTTTCGACCCCGCTTCGATCTCTCCTTTTATTCAGCCCTCCTCTCTTCTTAGGC 2039
Db 1342 CTTCTCTGTTTCGACCCCGCTTCGATCTCTCCTTTTATTCAGCCCTCCTCTCTTCTTAGGC 1401
Qy 2040 GCCCCCATATGGCCATATGATCTTATATATGGGCAACCCCGCCCTTGTAACTTCCCT 2099
Db 1402 GCCGAGAT-----CT 1411
Qy 2100 GACCTCAGATGACAAAGATTCTAACAGCCCTCTCTCCTCAAGCTCACTTACAGGCTTCT 2159
Db 1412 CTGAGAGCGTTAACGAGTTTAAACGAGCGCGCGCAAGCT----- 1456
Qy 2160 ACTTAGTCCAGCAGAAAGTCTGGAGACCTCTGGCGGAGCTTCAAGAAACAACTGAGAC 2219
Db 1457 -----TGAGAAATCCGCCCTCTCCTCCCGCCCTTAAAGCTTACTGGCGAA 1506
Qy 2220 GACCGGTGTTACCTCACCTTACCGAGTCCGGGACACAGTGTGGTCCGCGACACAGAGA 2279
Db 1507 GCGCTTTGGAATAGGCGGTTGCGTGTCTATATGTTATTTTCCACCATATGCGCT 1566
Qy 2280 CTAGAACTTAGAACCTCGCTGGAAAGGACCTTACACAGTCTCTGCTGACCAACCCACCG 2339
Db 1567 CTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTTCGACGAGCATTCTTAGGG 1626
Qy 2340 CCCTCAAAGTAGCGGATCGCAGCTTTGGATACACGCGGCCACGTAAGGCTGCCGACC 2399
Db 1627 GTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGTGAATGTCTGTAAGGAAGCAGTTC 1686
Qy 2400 CCGGGGTGGACCATCTCTAGACTGCGATGGGATGGAGCTGTATCATCTCTTCTTGGT 2459
Db 1687 CTCTG---GAAGCTTCTTGAAGCAAAACAGCTGTGTAGCGACCTTTGCG----- 1733
Qy 2460 AGCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAG 2519
Db 1734 -----AGCAGCGGAACCCCGACCTGCGGACAGTGTCTGCGCCAAAGCAGCTG 1788
Qy 2520 CGCAGCGTGGGTGACAGAGTGACCATCACTGTAAAGCCAGTCAAGATGTTGGTACTTC 2579
Db 1789 TATAAGATACACCTGCAAGGCGGCAACCCAGTGCACGTTGTGAGTTGGATAGTTG 1848
Qy 2580 TGTAGCTGTTGACAGCAGAACCCAGGTAAAGGCTCCAAAGCTGTGATCTACTGACATC 2639
Db 1849 TGGAAAGAGTCA----- 1860
Qy 2640 CACCGGCACACTGTTGTGTCACAGATTTCAGCGGTAGCGGTAGCGGTACCGACTTCAC 2699
Db 1861 ----- 1860
Qy 2700 CTTCAACATCAGCAGCCTCCAGCCAGAGGACATCGCCACTTACTTCCAGCAATATAG 2759
Db 1861 -----AATGGCT 1867
Qy 2760 CTTCTATCGTCTGTTCCGCCAAGGACCAAGTGGAAATCAACAGAGTGTGCTCAGGATC 2819
Db 1868 CTCTCAAGCGTATTTCAACAAGGGGCTGAAGGATGC-----CCAGAAGTATACCCCAAT 1920
Qy 2820 GGGTGGATCCGGCTCTGCTGCTCAGGATCGGAGTCCCACTGTGTGAGAGCGGTGGAGG 2879
Db 1921 GTATGGATCTGATCTCGGGCTC-----GGTGCAATGCTTTTACATGTGTTAGT 1971
Qy 2880 TGTGTGCAACCTTGGCGGCTCCCTGCGCTGTCTCTGCTCGCATCTGCTTTCGATTTTCAC 2939
Db 1972 CGAGTTTAAAAAACGTTCTAGGC---CCCCGAAACACGGGACGCTGTTTTCTTTGAAA 2028
Qy 2940 CACATTTGGATAGTTGGGTGAGACAGGCACTTGGAAAGGCTTTGATGGATTTGGAGA 2999
Db 2029 AACACGATGATAATATGGCCACAACCAAGGCGGAGGAGCTGTTCAACGGGTGTGCCCA 2088
Qy 3000 AATTTCATCCAGATAGCAGTACGATTAACTATGCGCCGCTCTCTTAAAGGATAGATTACAT 3059
Db 2089 TCCTGGTTCGAGCTGG----- 2103

QY 3060 ATCGGAGACAAAGCGCAAGAACACATTTGTTCTGCAAAATGGACAGCTTGAGACCGCGAAGA 3119
 Db 2104 -----ACGGCGAGCTGAACGSCCAAGTTTCAGGGTCTCGGGGAGGG 2146
 QY 3120 CACCGGGGTCTATTTTGTGCAAGCCCTTTACTTTGGCTTCCCTCGTGTGTTGCTTTATTTGGGG 3179
 Db 2147 CGAGGCGCATGCGCACCTACGCGCAAGCTGACCCCTGAAGTTTCATCTGCACACCGCGGAAGCT 2206
 QY 3180 CCAAGGACCCCGGTACCGTCTCCAGTCTTAAGCCCAACGACGCGCGCGCGAC 3239
 Db 2207 GCCCGTGCCTTGCCCAACCT-----CGTGACC 2234
 QY 3240 ACCAACACCGCGGCCCAACCATCGCGTGGAGCCCTCTCCCTCGCGCCACAGCGCGCTCG 3299
 Db 2235 ACCCTGACCTACGGGTGAGTGTCT--CAGCGCTACCCCGACACATGAAGACGACG 2292
 QY 3300 GCCAGCGCGGGGGCGAGTGACACACGAGGGGGTGGACTTTGGCCCTGGATCCCAAACT 3359
 Db 2293 AC-----TTCTTCAAGTCGCGCATGCCCGAAGG 2320
 QY 3360 CTGCTACCTGCTGGATGGAATCTCTTCACTATGTTGTCATTTCTCACTGCTGTTGTTCT 3419
 Db 2321 CTACGTCGAGGACGCGCACCATCTTCTTCAAGGACGACGCAACTACAAGACCGCGCGCA 2380
 QY 3420 GAGAGTGAAGTTTCAGCAGGAGCGCAGAGCCCGCGGTACACGAGGGGCCAGAACAGCT 3479
 Db 2381 GGTGAAGTTGAGGGGGACACCTGTTGAACCGCATCGAGCTGAAGGGCATCGACTTCA 2440
 QY 3480 CTATAACGAGCTCAATCTTAGGACGAGAGGAGTACGATGTTTGGACAAGAGAGCTGG 3539
 Db 2441 GGAGGACGCGCAACATCTCGGGCACAAAGCTGGAGTACAACACTACAACAGCCACAACTCTA 2500
 QY 3540 CCGGACCTGAGATGGGGGAAAGCCGAGAGAGAAACCTCAGGAAGGCTGTACAA 3599
 Db 2501 TATCATGGCCGACAAGCAAGAACGGCATCAAGCGCAACTTCAAGATCCGCCACAACAT 2560
 QY 3600 TGAACCTGCAGAAAGATAAGATGCGGAGGCTTACAGTGAGATTGGGATGAAGGGCGAGCG 3659
 Db 2561 CGAGGACGCGGTGCACCTGCGGACCACTACCGAGAACACCCCATCGGCGAGCG 2620
 QY 3660 CCGGAGGGCAAGGGGACGATGGCTTTACAGGGTCTCAGTACAGCCACCAAGGACAC 3719
 Db 2621 CCGCGTGTCTGCGCCGACAAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAAGACC 2680
 QY 3720 CTACGAGCCCTTACATCGAGCCCTGCCCCCTCGTAACTCGAGCGGCGCGGATCC 3779
 Db 2681 CAAACGAGAACCGCATCACATGGTCTGCTGGAGTTCTGTGACCGCGCGCGGATCACTCA 2740
 QY 3780 GGATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGCTCTAGTTTGTACTCA 3839
 Db 2741 CGGCATGGACGAGCTGTACAGTAAGTGCACCTGCAGCCCAAGCTT----- 2785
 QY 3840 CAATATCACAGCTGAAGCTATAGAGTACGAGCCATAGATAAAATAAAGATTTTATTT 3899
 Db 2786 -----ATCGATAAAATAAAGATTTTATTT 2810
 QY 3900 AGTCTCCAGAAAAGGGGGGATGAAGACCCCACTGTAGTTTGGGAGCTAGCTTAA 3959
 Db 2811 AGTCTCCAGAAAAGGGGGGATGAAGACCCCACTGTAGTTTGGCAAGCTAGCTTAA 2870
 QY 3960 GTAACGCCATTTTGAAGCATGGAATAATACATACTGAGAATAGAGAAGTTTCAGATCAA 4019
 Db 2871 GTAAACCCATTTTGCAGGCGATGGAATAATACATACTGAGAATAGAGAAGTTTCAGATCAA 2930
 QY 4020 GGTTAGGAAACAGAGACACAGCAATATGGGGCAACAGGATATCTGTGTGAAGCAGTTTC 4079
 Db 2931 GGTTAGGAAACAGAGACAGC----- 2951
 QY 4080 CTGCCCCCTCAGGGCCAGAAACAGTTGGAAACAGGAGATATGGGCAACAGGATATCT 4139
 Db 2952 -----AGAAATATGGGCAACAGGATATCT 2976
 QY 4140 GTGGTAAGCAGTTCTCTGCCCCCGCTCAGGGCCAAAGAACAGATGGTCCCCCAGATCGGTC 4199

Db 2977 GTGGTAAGCAGTTCTCTGCCCCGGCTCAGGGCCAAAGACAGATGGTCTCCAGATCGGCTCC 3036
 QY 4200 CGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGACCTG-AA 4258
 Db 3037 CGCCCTCAGCAGTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCAAGACCTGAAA 3096
 QY 4259 ATGACCTGTGCTTTATTTGAACTAAACAACTCAGTTTCTGCTTCTGCTTCTGTTGCGCGC 4318
 Db 3097 ATGACCTGTGCTTTATTTGAACCAATCAGTTTCTGCTTCTGCTTCTGTTTGGCGC 3156
 QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAGTCTCTCC 4378
 Db 3157 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCGCAGTCTCTCC 3216
 QY 4379 GATAGACTGCTGCGCGCGGTACCCGTGTTCTCAATAAACCCCTCTTGCAAGTTGCATCCA 4438
 Db 3217 GATAGACTGCTGCGCGCGGTACCCGTGTTCTCAATAAACCCCTCTTGCAAGTTGCATCCA 3276
 QY 4439 CTGCTGCTCGCTTCTGTTGGAGGGTCT-CTCTGAGTGATTGACTACCCGTGAGCGG 4497
 Db 3277 CTGCTGCTCGCTTCTGTTGGAGGGTCTCTCTGAGTGATTGACTACCCGTGAGCGG 3336
 QY 4498 GGTCTTTCAGTTTCTCCACCTACACAGGTCTCACTAAACATTCCTGATGTGCGCGCAGGGA 4557
 Db 3337 GGGTCTT----- 3343
 QY 4558 CTCCTCAGCCCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617
 Db 3344 ----- 3343
 QY 4618 ACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCC 4677
 Db 3344 ----- 3343
 QY 4678 ACACGACTTTTAAAGATTTTATGCTCTCTGATGAGGGAATTTAGTCAATCTATCCTCGT 4737
 Db 3344 ----- 3343
 QY 4738 CTATTTTGTGCTGCTTCTCGGTATTTTAAATTTCTAGTTTGCACTCCCTTCTGAGAGCAC 4797
 Db 3344 ----- 3343
 QY 4798 GCGGATTGCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAAAGGTTGCTGGGCT 4857
 Db 3344 ----- 3343
 QY 4858 CAGTGTGAGATTTTGCATATAAAGGGGTCTCTGCCCCCTGTGTACAGACAGATCGGAATCT 4917
 Db 3344 ----- 3343
 QY 4918 AGAGTGCATATCTCAGAGTCCCCCGGGTCTCGGGGGCTCTGATCTCAGGGCATCTTTGCTTA 4977
 Db 3344 ----- 3343
 QY 4978 GAGATCCTTACGCCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGTAATCATGCT 5037
 Db 3344 -----TCAGTATTCGTAATCATGCT 3363
 QY 5038 CATAGCTGTTTCTGTTGAAATTTATCCGCTCACAATTCACACAACATACGAGCGG 5097
 Db 3364 CATAGCTGTTTCTGTTGAAATTTATCCGCTCACAATTCACACAACATACGAGCGG 3423
 QY 5098 GAAGCATAAAGTGAAGCTCGGGTGCCTAATGATGAGTGAGTCACTCAATTAATTTGCGT 5157
 Db 3424 GAAGCATAAAGTGAAGCTCGGGTGCCTAATGATGAGTGAGTCACTCAATTAATTTGCGT 3483
 QY 5158 TCGGCTCACTGCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGCAATTAATTAATCG 5217
 Db 3484 TCGGCTCACTGCGCTTCCAGTCGGGAAACCTGTCGTCAGCTGCAATTAATTAATCG 3543
 QY 5218 GCCAACGCGCGGAGAGCGGTTTGGCTATTTGGCGCTCTTTCCGCTTCTCTCGCTCACTG 5277

Db 3544 GCCAACGGCGGGGAGAGCGGTTTGGCTATTGGCGGCTCTTCCGCTCTCTCGCTCACTG 3603
Qy 5278 ACTCGCTCGCTCGGTCGTTGCGGTGCGGCGAGCGGTATCAGCTCACTCAAGCGGTAA 5337
Db 3604 ACTCGCTCGCTCGGTCGTTGCGGTGCGGCGAGCGGTATCAGCTCACTCAAGCGGTAA 3663
Qy 5338 TACGGTTATCCACAGATCAGGGATTAACGAGAAAGAACATGTGAGCAAAAGCGCAGC 5397
Db 3664 TACGGTTATCCACAGATCAGGGATTAACGAGAAAGAACATGTGAGCAAAAGCGCAGC 3723
Qy 5398 AAAAGGCGAGAACCGTTAAAGGCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCC 5457
Db 3724 AAAAGGCGAGAACCGTTAAAGGCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCC 3783
Qy 5458 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACAGCACTAT 5517
Db 3784 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGAAACCCGACAGCACTAT 3843
Qy 5518 AAAGATACGAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGC 5577
Db 3844 AAAGATACGAGCGGTTTCCCGCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCCCTGC 3903
Qy 5578 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT 5637
Db 3904 CGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCT 3963
Qy 5638 CACCGCTAGGATCTCAGTTCCGTTGAGTTCGCTTCCGAGCTCGGCTGTGTGCAACG 5697
Db 3964 CACCGCTAGGATCTCAGTTCCGTTGAGTTCGCTTCCGAGCTCGGCTGTGTGCAACG 4023
Qy 5698 AACCCCGGTTACGCCGACCGCTGCGCTTATCCGTTAATCCGTTAATCCGTTGAGTCCAAAC 5757
Db 4024 AACCCCGGTTACGCCGACCGCTGCGCTTATCCGTTAATCCGTTAATCCGTTGAGTCCAAAC 4083
Qy 5758 CGGTAAGACACGACTTATCCGCTACGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAGAA 5817
Db 4084 CGGTAAGACACGACTTATCCGCTACGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAGAA 4143
Qy 5818 GGTAATGAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAGAA 5877
Db 4144 GGTAATGAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACTACGGCTACACTAGAA 4203
Qy 5878 GGACAGATTTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGGTA 5937
Db 4204 GGACAGATTTGGTATCTGGCTCTGTGAAGCCAGTTACCTTCGGAAGAAAGAGTTGGTA 4263
Qy 5938 GCTCTTGATCCGGCAAAACACCGCTGTGTAGCGGTGTTTTTTTGTGTAAGCAGC 5997
Db 4264 GCTCTTGATCCGGCAAAACACCGCTGTGTAGCGGTGTTTTTTTGTGTAAGCAGC 4323
Qy 5998 AGATTACGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGGTCTG 6057
Db 4324 AGATTACGCGAGAAAAAGGATCTCAAGAGATCCTTTGATCTTTTCTACGGGGTCTG 4383
Qy 6058 ACCTCAGTGAACGAAAACTCAGCTTAAAGGATTTTGGTCAAGATTAATCAAAAAAGGA 6117
Db 4384 ACCTCAGTGAACGAAAACTCAGCTTAAAGGATTTTGGTCAAGATTAATCAAAAAAGGA 4443
Qy 6118 TCTTACCTAGATCCTTTTAAATTAATAAAGATTTTAAATCAATCAATCAATCAATCAATCAAT 6177
Db 4444 TCTTACCTAGATCCTTTTAAATTAATAAAGATTTTAAATCAATCAATCAATCAATCAATCAAT 4503
Qy 6178 AGTAACTTGGTTCAGACATTTACCAATGCTTAATCAGTGGGACCTTATCTCAGCGATCT 6237
Db 4504 AGTAACTTGGTTCAGACATTTACCAATGCTTAATCAGTGGGACCTTATCTCAGCGATCT 4563
Qy 6238 GTCTATTTTGGTTCATCAATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGG 6297
Db 4564 GTCTATTTTGGTTCATCAATAGTTGCTGACTCCCGCTCGTGTAGATAACTACGATACGGG 4623
Qy 6298 AGGCTTACCATCTGGGCCCAAGTGTGCAATGATACCGCGAGACCCAGCTCACCAGGCTC 6357
Db 4624 AGGCTTACCATCTGGGCCCAAGTGTGCAATGATACCGCGAGACCCAGCTCACCAGGCTC 4683

Qy 6358 CAGATTTATCAGCAATAAAACAGCCAGCCGAGAGCGGCGAGAGTGGTCTCTGCAA 6417
Db 4684 CAGATTTATCAGCAATAAAACAGCCAGCCGAGAGCGGCGAGAGTGGTCTCTGCAA 4743
Qy 6418 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAGCTAGAGTAAGTATTCGC 6477
Db 4744 CTTTATCCGCTCCATCCAGTCTATTAAATTTGTTGCGGGAGCTAGAGTAAGTATTCGC 4803
Qy 6478 CAGTTAATAGTTTGGCGAAAGTTGTTGCTCATTTGTCAGAGC - TCGTGGTGTCAAGCTCGT 6536
Db 4804 CAGTTAATAGTTTGGCGAAAGTTGTTGCTCATTTGTCAGAGATCGTGGTGTCAAGCTCGT 4863
Qy 6537 CGTTTGGTATGGCTTCAATCAGCTCCGTTTCCCAACGATCAAGCGAGTTCATCATGATCCC 6596
Db 4864 CGTTTGGTATGGCTTCAATCAGCTCCGTTTCCCAACGATCAAGCGAGTTCATCATGATCCC 4923
Qy 6597 CCATGTTGTCAGAAAAAGCGGTTAGCTCTTCGCTCCCGATCGTTGTGAGAGTAAGT 6656
Db 4924 CCATGTTGTCAGAAAAAGCGGTTAGCTCTTCGCTCCCGATCGTTGTGAGAGTAAGT 4983
Qy 6657 TGGCCGAGGTTTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCTATGC 6716
Db 4984 TGGCCGAGGTTTATCACTCATGTTATGGCAGCACTGCATAATCTCTTACTGTCTATGC 5043
Qy 6717 CATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCAATCTTGAGAAAGT 6776
Db 5044 CATCCGTAAGATGCTTTTCTGTGATCGGTGAGTACTCAACCAAGTCAATCTTGAGAAAGT 5103
Qy 6777 GTATGCGGCGACGAGTTGCTCTTCCCGGGGTCAATAACGGGATAATACCGCGCACATA 6836
Db 5104 GTATGCGGCGACGAGTTGCTCTTCCCGGGGTCAATAACGGGATAATACCGCGCACATA 5163
Qy 6837 GCAGAACTTTAAAGTCTCATCATTCGAAAAAGCTTCTCGGGCGGAAAACTCTCAAGGA 6896
Db 5164 GCAGAACTTTAAAGTCTCATCATTCGAAAAAGCTTCTCGGGCGGAAAACTCTCAAGGA 5223
Qy 6897 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAACCACTGATCTTTCAG 6956
Db 5224 TCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGCAACCACTGATCTTTCAG 5283
Qy 6957 CATCTTTTATCTTTCACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATCCCGCAA 7016
Db 5284 CATCTTTTATCTTTCACAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATCCCGCAA 5343
Qy 7017 AAAAGGGAATAAGGGGACACGGAATGTTGAATCTACTACTCTCTCTTTTCAATATT 7076
Db 5344 AAAAGGGAATAAGGGGACACGGAATGTTGAATCTACTACTCTCTCTTTTCAATATT 5403
Qy 7077 ATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGA 7136
Db 5404 ATTGAAGCATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTGAATGTATTTAGA 5463
Qy 7137 AAAATAAACAAATAGGGTTTCCGCGCATTTTCCCGAAAGTGCCACCTGACGCTTAAG 7196
Db 5464 AAAATAAACAAATAGGGTTTCCGCGCATTTTCCCGAAAGTGCCACCTGACGCTTAAG 5523
Qy 7197 AAACCATTTATCATGACATTAACCTTAAATAGCGGTATCACGAGGCCCTTTTCGTC 7256
Db 5524 AAACCATTTATCATGACATTAACCTTAAATAGCGGTATCACGAGGCCCTTTTCGTC 5583
Qy 7257 TCGCGGCTTTCGCTGATGAGCGGTGAAAACTCTTGACATCATGACGCTCCCGAGAGCGGTCA 7316
Db 5584 TCGCGGCTTTCGCTGATGAGCGGTGAAAACTCTTGACATCATGACGCTCCCGAGAGCGGTCA 5643
Qy 7317 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGAACGCCGCTCAGGGCGGCTCAGCGGCTG 7376
Db 5644 CAGCTTGTCTGTAAGCGGATGCGGGAGCAGAACGCCGCTCAGGGCGGCTCAGCGGCTG 5703
Qy 7377 TTGGCGGCTTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGTCTGAGAGTGC 7436
Db 5704 TTGGCGGCTTCCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTTGTAAGTCTGAGAGTGC 5763

Db 1642 CGGAATTCGTTAACTCGAGGATCCACTAGTAACCGCGCCGACGTGTGCTGGAAATTAATTCG 1701
 Qy 2102 C--CCTGACATGACAAGAGTTACTAAACAGCCCTCTCTCCAAGCTCACTTACAGGCTTCT 2159
 Db 1702 CTGCTCGAGGGCGCGTGTGGGGTGAGTACTCCCTCTCAAAGCGGGCATGACTTCT 1761
 Qy 2160 ACTTAGTCCAGCAGAACTCTGGAGACCTCTGGCGGCAGCCTTACCAAGAACAACTGGACC 2219
 Db 1762 CGCGTTAAGATTGTCAAGTTTCCAAAACGAGGAGGATTTGATATTCACTGCGCCGCGGTG 1821
 Qy 2220 GACCGGTGTACTTACCTTACCGAGTCGGGACACAGTGTGGTCCGCGACACACAGA 2279
 Db 1822 ATGCTTTGAGGGTGGCGCGCTCCATCTGGTCAGAAAAGACAATCTTTTGTGTCAAGC 1881
 Qy 2280 CTAAGAACCTAGAACCTCGCTGGAAAGACCTTACACAGTCTCTGTCCACCAACCCCAACCG 2339
 Db 1882 TTGAGGTGTGCGAGCTTGAGATCTGGCCATACACTTGAAGTACATGACATCACTTTG 1941
 Qy 2340 CCCTCAAAGTAGCGGCATCGAGCTTGGATACACGCCGCCACGTGAAGGCTGCCGACC 2399
 Db 1942 CTTTCTCTCACAGGTGTCCACTCCCAAGTCCAACTCGAGTCCGATCGATGATCT 2001
 Qy 2400 CCGGGGTGGACCATCTCTAGACTGCTCATGGGATGGAGCTGTATCATCTCTTCTTGGT 2459
 Db 2002 AGGCGGGCCAAATTCGCCCTCTCCCTCCCGCCCTTAACGTTTACTGGCCGGAAGCGCTT 2061
 Qy 2460 AGCAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCACAGCCCAAGCAGCTGAG 2519
 Db 2062 GGAATAGGCGGTGTGTGTGTCTATATGTGATTTTCCACCATTTTGGCGTCTTTGG 2121
 Qy 2520 CGCCAGCGTGGTGACAGAGTGACCATCACCTGTAAGGCCAGTCAAGATGTGGGTACTTTC 2579
 Db 2122 CAATGTGAGGCCCGGAAACCTGGCCCTGTCTCTTGAACGAGCATTCCTAGGGGTCTTTC 2181
 Qy 2580 TGTAGCTTGTGACMGAGAAAGCCAGGTAAAGCTTCCAAAGCTGTGATCTACTGACATC 2639
 Db 2182 -----CCCTCTCGCCAAAGAAATGCAAGGTCTGTGTAATGTGTGTAAGGAAGCAGTTC 2234
 Qy 2640 CACCGGACACTGTGTGCCAAGCAGATTCAGCGGTAGCGGTAGCGGTACCGACTTCAC 2699
 Db 2235 CTCTGGAAGCTTCTTGAAGACAACAACAGCTCTGTAGCAGCCCTTTGACGAGCGGAACC 2294
 Qy 2700 CTTCAACATCAGAGCTCTCAGCCAGAGGACATCGCCACCTACTACTGCCAGCAATATAG 2759
 Db 2295 CCCCACCTGGCGACAGGTGCCTCTCGGSCCAAAGCCAGT-----GTATAGATACA 2347
 Qy 2760 CTTCTATCGGTCTTGGCCAAAGGACAAGGTGGAATCAAAAGAGTGTGCTCAGGATC 2819
 Db 2348 CTTGAAAGGGCGCACAAACCCAGTGCACCGCTTGTAGTTGGATAGTTGTGGAAGAGTTC 2407
 Qy 2820 GGGTGGATCCGGCTCTGTGTGCTCAGGATCGAGGTCCAACTGGTGGAGAGCGGTGGAG 2879
 Db 2408 AAATGGCTCTCTCAAGGTATGTAACAAGGGGTGAAGGATGCCCAAGAGGTACCCCAT 2467
 Qy 2880 TGTGTGCAACCTGGCGGTCCCTCGCGCTCTCTCTGCTCCGCACTCTGGCTTCGATTTTCA 2939
 Db 2468 TGTATGGGAATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTG-----TTTAG 2519
 Qy 2940 CACATATTGGATAGTTGGGTGAGACAGGCACTTGAAGAAAGGTCTTGAAGTGGATTTGAGA 2999
 Db 2520 TCGAGGTTAAAAAGCTCTAGGCCCCCGGAACCAACCGGGAGCGTGTGTTTCTCTTT---GA 2575
 Qy 3000 AATTTCATCCAGATGACGATAGTAACTATGCGCGCTCTCTAAGGATAGATTACAT 3059
 Db 2576 AAAACACGATGATAGCTTGCACAACCCCGGATAATTCTCGAGCCAAATATGGGATCG 2635
 Qy 3060 ATCGGAGACAAACGCAAGAACACATTTCTCTGCAATGACAGCCCTGACACCCGA--A 3117
 Db 2636 GCCATTGACAGATGGATTGCAAGGTTCTCGGCGCGCTTGGGTGGAGAGGCTATTTC 2695
 Qy 3118 GACACCGGGGTCTATTTTGTGCAAGCCTTTACTTTCGGCTTCCCTGCTTGTATTTGG 3177

Db 2696 GGCTATGATCGGCGACAAACAGACAATCGGTGCTCTGTATGTCGCGCGTGTTCGGGTGTCA 2755
 Qy 3178 GGCCAAAGGAGCCCGGTCAACGCTCTCCAGTGTCTAAGCCACACAGCAGCAGCGCGGA 3237
 Db 2756 GCGCAGGGCGCCCGGTCTTTTGTCAAGACCGACCTGTCCGTGCGC---TGAATGNA 2812
 Qy 3238 CCAACAACACCGGCGCCACCATCGGTTCGAGCCCTGTCTCTGCGCCCGCAGAGCGGT 3297
 Db 2813 CTGAGGACGAGGAGCGCGGTATCGTGTGCGCACGACGCGGTCTCTTTCGCGAGCT 2872
 Qy 3298 CGCCAGCGGCGGGGCGAGTGCACACGAGGGGCTGGACTTCGCCCTGGATCCCAA 3357
 Db 2873 GTGCTCGACGTTGTCACTGAAGCGGAAGGACTGGCTGCTATTGGGCGAATATGCGGG 2932
 Qy 3358 CTCTGCTACCTGCTGGATGGAATCTCTTATCTATGCTGTCATTTCTCACTGCTCTGTTTC 3417
 Db 2933 CAGGATCTCTGTCTCATCTCACCTTGTCTCTCCGAGAAAGTATCCATCATGGCTGATGCA 2992
 Qy 3418 CTGAGAGTGAAGTTTACGAGGAGCGAGAGCCCGCGGTACCCAGAGGGCCAGAACAG 3477
 Db 2993 ATGCGGGGCTGCATACGCTTGTATCGGCTACTGCCCCATTCGACCACCAAGCGAACA 3052
 Qy 3478 CTCTATACGAGCTCAATCTAGGACGAGAGAGGAGTACGATGTTTTGGACAAGAGAGT 3537
 Db 3053 CGC-----ATCGAGCGAGCACGTACTCGGATGGAAGCGGCTTTGTGATCAGGAT 3103
 Qy 3538 GGCCGGACCTGAGATGGGGGAAAGCCGAGAGGAAGAACCCCTCAGGAAGGCTGTAC 3597
 Db 3104 GATCTGAGCGNAGAGCATCAGGG--GCTCGCGCCAGCGGAACTGTTGCCAGGCTCAAG 3161
 Qy 3598 AATGAATGCAAGAAGATTAAGATGCGGAGCGCTTACAGTGAAGATTTGGGATGAAGGCGAG 3657
 Db 3162 CGCGCATGCCCGAGCGGAGGATCTCGTCTGTGACCCATGGCGATCGCTGTTGCCGAATA 3221
 Qy 3658 CGCCGAGGGGCAAG--GGCACGATGGCTTTACAGGGTCTCAGTACAGCCACCAAG 3715
 Db 3222 TCATGSGTGGAAATGCGCGCTTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGGCGG 3281
 Qy 3716 ACACCTACGAGCGCTTTCACATGACGCGCTGCCCCCTCGCTAA-----CTCGACGCGCC 3771
 Db 3282 ACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTTCTGAAGAGTTTGGCGCGAAT 3341
 Qy 3772 GCGGATCCGGATTAAGTCCAAATTTGTTAAAGACAGGATATCAAGTGTGTCAGGCTCTAGTTT 3831
 Db 3342 GGGCTGACCGCTTCTCTGCTTTACGCTATCGCGCTCCCGATTTGCGAGCGCATCGCT 3401
 Qy 3832 TGACTCAACATATACACGCTGAAGCTATAGAGTAGAGCCATAGATATAAATAAGA 3891
 Db 3402 TCTATCGCTTCTTTGACGAGTTCTTCTGAGCGGACTCTGGGGTTTCGATAAATAAAGA 3461
 Qy 3892 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGC 3951
 Db 3462 TTTTATTTAGTCTCCAGAAAAGGGGGAATGAAGACCCACCTGTAGGTTTGGCAAGC 3521
 Qy 3952 TAGCTTAAGTAACCCCAATTTTGAAGGCATGG-AAAAATACATAAATCTGAGAATAGAGAAGT 4010
 Db 3522 TAGCTTAAGTAACCCCAATTTTGAAGGCATGGAATAAATACATACTGAGAATAGAGAAGT 3581
 Qy 4011 TCAGATCAAGGTTAGGAACAGA-GAGACAGAGAAATATGGGCCAAAAGGATATCTGTGG 4069
 Db 3582 TCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAAAGGATATCTGTGG 3641
 Qy 4070 TAACAGTTCTCTGCCCCC-GCTCAGGGCCAAAGACAGTTGGAACAGGAGAAATATGGCCAA 4128
 Db 3642 TAACAGTTCTCTGCCCCCAGGCTCAGGGCCAAAGACAGATGGAACAGTGAATATGGGCCAA 3701
 Qy 4129 ACAGGATATCTGCTGAAGCAGTTCTCTGCCCGGCTCAGGGCCAAAGACAGATGCTGCC 4188
 Db 3702 ACAGGATATCTGCTGAAGCAGTTCTCTGCCCGGCTCAGGGCCAAAGACAGATGCTGCC 3761
 Qy 4189 AGATGCGGTCCCGCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCA 4248
 Db 3762 AGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGGCCCA 3821

QY 4249 AGGACCTGAAATGACCCCTGTGCTTATTTTGAATCAACCAATCAGTTTCGCTTCTCGCTTCT 4308
DB 3822 AGGACCTGAAATGACCCCTGTGCTTATTTTGAATCAACCAATCAGTTTCGCTTCTCGCTTCT 3881
QY 4309 GTTTCGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGG 4368
DB 3882 GTTTCGCGCTTCTGTCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGCGCGG 3941
QY 4369 CCAGTCTCCGATAGACTCGCTGCGCGGGTACCGGTGTTCTCAATAAACCCTCTTGCAG 4428
DB 3942 CCAGTCTCCGATAGACTGAGTGCAGCGCGGTACCGGTGTTCTCAATAAACCCTCTTGCAG 4001
QY 4429 TTGCAATCCGACTCGTGTCTCGCTGCTTCTCGGAGGGTCT-CTCTGAGTGATGACTAC 4487
DB 4002 TTGCAATCCGACTGAGTGTCTCGCTGTTCTCGGAGGGTCTCTCTGAGTGATGACTAC 4061
QY 4488 CCGTCAGCGGGTCTTTTCACTTCTCCCACTACACAGCTCTCACTAAATTCCTGATGT 4547
DB 4062 CCGTCAGCGGGTCTTTTCACTTTCGGGGTCTCTCGGGATCGGAGACCCCTGCCAGGG 4121
QY 4548 GCGGAGGAGCTCCGTGACCCGGTTTTGTGTTTATAATAAATGCAAGAACAGTGTCCC 4607
DB 4122 ACCACGACCCACCGGAGGTAAAGCTG----- 4151
QY 4608 TTCAAGCCAGACTACATCTGACTCTCGCTTTATAAAGAAATGTTGAAGGCTCTGTGG 4667
DB 4152 -----GCTGCTCGCGGTTTCGGTGATGACGGTGAACCTCTG-AC 4193
QY 4668 ACTATCTGCACACGACTTTTTAAGATTTTATATGCTCTCGATGAGGATTTAGTCAAT 4727
DB 4194 ACATGAGCTCCCGGAGACGGTCAAGCTTGTCTGTAAGCGATGCCGGAGCAGACAAG 4253
QY 4728 CTATCTCTGTATTTTGTGCTGCTTCTCGTATTTTAAATTTCTAGTTTGCATCCCTTC 4787
DB 4254 CCGCTCAGGCGCGTCAGCGGGTGTGGG----- 4283
QY 4788 CTGAGACGACGGGATTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAGAAGT 4847
DB 4284 ----- 4283
QY 4848 TGCCTGGGCTCAGTGTGAGATTTTGCCATAAAGGGTCTCTGCCCTGTGTACAGACAG 4907
DB 4284 ----- 4283
QY 4908 ATCGGAATCTAGAGTGCACTACAGATCCCGGGTTCGGGGCTCTGATCTCAGGGCA 4967
DB 4284 -----GCTGTGCGGGCGCAGCCATGA----- 4304
QY 4968 TCTTTGCTAGAGATCCTCTACGCGGACGCATCGTGGCGGGTACCGAGCTCGAATTCG 5027
DB 4305 ----- 4304
QY 5028 TAATCATGTGTCATAGTGTTCCTGTGTGAATTTGTTATCCGCTCACAAATCCACACAAAC 5087
DB 4305 -----CCAGTTCAG 4314
QY 5088 ATACGAGCCGGAAGCAATAAGTGTAAAGCTGGAAGCTGGGTGCTTAATGAGTGAGTAACTACA 5147
DB 4315 TAGCGATAGCGGAGTGATACTGGGTTAACTATGCGGCATCAGAGCAGATTTGACTGAGA 4374
QY 5148 TTAATTTGCGTTCGCTCACTGCGCGCTTTCAGTCCGGAAACCTGTGCGGAGCTGCAT 5207
DB 4375 GT-----GCACCATATGCGG 4389
QY 5208 TAATGAATCGGCAACGCGCGGGAGAGCGGTTTTCGTTATTTGGGCGCTCTTCCGCTTCC 5267
DB 4390 TGTGAATACCGCACAGATGCGTAAGGAGAAATACCGCATCAGGGCTCTTCCGCTTCC 4449
QY 5268 TCGCTCACTGACTCGTTCGCTCGGTGCTGTTTCGGCTGCGGAGCGGTATCAGCTCACTCA 5327
DB 4450 TCGCTCACTGACTCGTTCGCTCGGTGCTGTTTCGGCTGCGGAGCGGTATCAGCTCACTCA 4509

QY 5328 AAGCGGTAAATACGGTTATCCACAGATCAGGGATAACCGAGGAAAGAAACAATGTGAGCA 5387
DB 4510 AAGCGGTAAATACGGTTATCCACAGATCAGGGATAACCGAGGAAAGAAACAATGTGAGCA 4569
QY 5388 AAGGCCAGCAAAAGGCCAGGAAACGTTAAAGGCGCGTGTCTGGCGTTTTTCCATAGG 5447
DB 4570 AAGGCCAGCAAAAGGCCAGGAAACGTTAAAGGCGCGTGTCTGGCGTTTTTCCATAGG 4629
QY 5448 CTCGCGCCCTTGAGCAGCATCAAAATCGACCTCAAGTCAGAGGTGGGAAACCCG 5507
DB 4630 CTCGCGCCCTTGAGCAGCATCAAAATCGACCTCAAGTCAGAGGTGGGAAACCCG 4689
QY 5508 ACAGACTATTAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTT 5567
DB 4690 ACAGACTATTAAGATACACAGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTGTT 4749
QY 5568 CCGACCTTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCTT 5627
DB 4750 CCGACCTTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCTT 4809
QY 5628 TCTCATAGTCACTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTCGAAGCTGGC 5687
DB 4810 TCTCATAGTCACTGTAGTATCTCAGTTTCGGTGTAGTTCGTTCCGTCGAAGCTGGC 4869
QY 5688 TGTGTGACAGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGCTAACTATCTCTT 5747
DB 4870 TGTGTGACAGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGCTAACTATCTCTT 4929
QY 5748 GAGTCCAAACCCGGTAAGACAGACTTATCGCCACTTCGGCAGCAGCACCTGGTAACAGGAT 5807
DB 4930 GAGTCCAAACCCGGTAAGACAGACTTATCGCCACTTCGGCAGCAGCACCTGGTAACAGGAT 4989
QY 5808 AGCAGCAGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACGCG 5867
DB 4990 AGCAGCAGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGCTAACTACGCG 5049
QY 5868 TACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATA 5927
DB 5050 TACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATA 5109
QY 5928 AGAGTTGTAGTCTTTGATTCGGGCAAAACAAACCCCGCTGTGAGCGGTGGTTTTTGT 5987
DB 5110 AGAGTTGTAGTCTTTGATTCGGGCAAAACAAACCCCGCTGTGAGCGGTGGTTTTTGT 5169
QY 5988 TGCAGCAGCAGATTTACCGGCAAAAGATCTCAAGAGATCTTGTATCTTTCT 6047
DB 5170 TGCAGCAGCAGATTTACCGGCAAAAGATCTCAAGAGATCTTGTATCTTTCT 5229
QY 6048 ACGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTA 6107
DB 5230 ACGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGATTTTGGTCATGAGATTA 5289
QY 6108 TCAAAAAGGATTTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 6167
DB 5290 TCAAAAAGGATTTTCACTTAGATCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 5349
QY 6168 AGTATATAGTAACTTGTCTCAGAGTTACCAATGCTTAATCAGTGAGGACCTATC 6227
DB 5350 AGTATATAGTAACTTGTCTCAGAGTTACCAATGCTTAATCAGTGAGGACCTATC 5409
QY 6228 TCAGGATCTGTCTATTTTCTGTTTCTGCTGACTCCCGTTCGTTGAGTAACT 6287
DB 5410 TCAGGATCTGTCTATTTTCTGTTTCTGTTTCTGCTGACTCCCGTTCGTTGAGTAACT 5469
QY 6288 AGCATACGGGAGGTTCACATCTGCGCCCTAGTGTGCAATGATACCGGAGACCCACG 6347
DB 5470 AGCATACGGGAGGTTCACATCTGCGCCCTAGTGTGCAATGATACCGGAGACCCACG 5529
QY 6348 TCACCGCTCCAGATTTATCAGCAATAACCGAGCGGAGGCGGAGCGGAGGAGT 6407
DB 5530 TCACCGCTCCAGATTTATCAGCAATAACCGAGCGGAGGCGGAGGCGGAGGAGT 5589
QY 6408 GGTCTCAGAACTTTATTCGCGCTCCATCCAGTCTATTAATTTGTTGCGGGAAGCTAGATTA 6467

```

Db 5590 GGTCTGCACTTTATCGCCTCCATCCAGTCTATTAATTTGTCGGGAAGCTAGAGTA 5649
Qy 6468 AGTAGTTCGCCAGTAAATAGTTTGGCAACGTTGTGTCATTTGCTACAGGC-TGTTGGTG 6526
Db 5650 AGTAGTTCGCCAGTAAATAGTTTGGCAACGTTGTGTCATTTGCTGCAAGGCATCGTGGTG 5709
Qy 6527 TCAGCTCGTGTGGTATGGCTTCATTCAGTCCGGTTCCTCAACGATCAAGGGAGTT 6586
Db 5710 TCAGCTCGTGTGGTATGGCTTCATTCAGTCCGGTTCCTCAACGATCAAGGGAGTT 5769
Qy 6587 ACATGATCCCATCTGTCGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTC 6646
Db 5770 ACATGATCCCATCTGTCGCAAAAAGCGTTAGTCTCTTCGGTCTCCGATCGTTGTC 5829
Qy 6647 AGAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTT 6706
Db 5830 AGAAGTAAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAATTTCTTT 5889
Qy 6707 ACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTGATCACTCAACCAAGTCAATC 6766
Db 5890 ACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGTGTGATCACTCAACCAAGTCAATC 5949
Qy 6767 TGAGAAATAGTGTATGCGCGGACCGAGTTGCTTTCGCCGGCTCAATACGGGATAATACC 6826
Db 5950 TGAGAAATAGTGTATGCGCGGACCGAGTTGCTTTCGCCGGCTCAATACGGGATAATACC 6909
Qy 6827 GCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGCGGAAA 6886
Db 6010 GCGCCACATAGCAGAACTTTAAAGTGTCTCATTTGGAACACGTTCTTCGGGCGGAAA 6069
Qy 6887 CTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCATCTGTCGCCAAC 6946
Db 6070 CTCTCAAGGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCATCTGTCGCCAAC 6129
Qy 6947 TGATCTCAGCATCTTTTACTTTTCCACAGCGTTTTCGGGTGACGAAACAGGAGGCAA 7006
Db 6130 TGATCTCAGCATCTTTTACTTTTCCACAGCGTTTTCGGGTGACGAAACAGGAGGCAA 6189
Qy 7007 AATGCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTT 7066
Db 6190 AATGCGCAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACTCATACTCTTCTCTT 6249
Qy 7067 TTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA 7126
Db 6250 TTTCAATATATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA 6309
Qy 7127 TGTATTTAGAAAATAAACAATAGGAGTTTCGGGTCACATTAATAAATAGTGCACCT 7186
Db 6310 TGTATTTAGAAAATAAACAATAGGAGTTTCGGGTCACATTAATAAATAGTGCACCT 6369
Qy 7187 GACGCTTAAGAAACCATTTATCATGACATTAACCTATAAATAAGGCGTATCAGGAG 7246
Db 6370 GACGCTTAAGAAACCATTTATCATGACATTAACCTATAAATAAGGCGTATCAGGAG 6429
Qy 7247 CCCTTTCGTCT 7257
Db 6430 CCCTTTCGTCT 6440

```

RESULT 11
 US-09-963-206B-4
 ; Sequence 4, Application US/09963206B
 ; Patent No. US20020123076A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ferrick, David A.
 ; APPLICANT: Swift, Susan E.
 ; APPLICANT: Armstrong, Randall
 ; APPLICANT: Fox, Bryan
 ; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
 ; TITLE OF INVENTION: Secretion and Switch Rearrangement
 ; FILE REFERENCE: A-66038-3/RMS/JUD/DLR
 ; CURRENT APPLICATION NUMBER: US/09/963.206B

; CURRENT FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US 09/076,624
 ; PRIOR FILING DATE: 1998-05-12
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 6219
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic
 US-09-963-206B-4

Query Match 40.3%; Score 3082.2; DB 9; Length 6219;
 Best Local Similarity 73.7%; Pred. No. 0;
 Matches 4587; Conservative 0; Mismatches 958; Indels 680; Gaps 25;

```

Qy 958 CCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATAGACTG 1017
Db 662 CAGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGCGCCAGTCTCCGATAGACTG 721
Qy 1018 CGTCGCGGGGTACCCGTTATCCCAATAAAGCCCTCTTGTCTTTGCAATCCGAATCGTGA 1077
Db 722 AGTCGCGGGGTACCCGTTATCCCAATAAAGCCCTCTTGTGAG-TTGCATCCGATGTGCT 780
Qy 1078 CTGCTGATCTTGGGAGGCTCTCTCAGATTGATTGATGCTGCC-ACCTCGGGGGTCTTT 1136
Db 781 CTGCTGTTCTTGGGAGGCTCTCTCTGAGTGTGACTACCCGTCAAGCGGGGTCTTT 840
Qy 1137 CATTTGGAGGTTCCACCGAGATTTGGAGACCCCTCCCGAGGACCCAGCCGCCGCC 1196
Db 841 CATTTGGGGGCTCGTCGGGATCGGGAGACCCCTGCCAGGGACCCCGA-CCCAACCAC 899
Qy 1197 GGGAGTAAAGCTGCCAGCAACTTATCTGTCTGTCTGCGATTGTCTAGTGTCTATGACTG 1256
Db 900 GGGAGTAAAGCTGCCAGCAACTTATCTGTCTGTCTGCGATTGTCTAGTGTCTATGACTG 959
Qy 1257 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCGTG 1316
Db 960 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCGTG 1019
Qy 1317 GTGAACTGACGAGTTCCGGAACACCCCGCCGCAACCTTGGAGAGCTCCAGGAGCTTCG 1376
Db 1020 GTGAACTGACGAGTTCCGGAACACCCCGCCGCAACCTTGGAGAGCTCCAGGAGCTTCG 1079
Qy 1377 GGGCCCGTTTTTGGGCCCGACCTGAGTCTTAAATCCCGATCCGATCTTTAGGACTCTTGGT 1436
Db 1080 GGGCCCGTTTTTGGGCCCGACCTGAGTCTTAAATCCCGATCCGATCTTTAGGACTCTTGGT 1139
Qy 1437 GCACCCCTTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTC 1496
Db 1140 GCACCCCTTTAGAGGAGGATATGTGGTTCTGGTAGGAGACGAGAACCTTAAACAGTTC 1199
Qy 1497 CCGCTCGCTCTGAATTTTGTCTTGGAGCCGAGCCGCGCGCGCTTGTGTC 1556
Db 1200 CCGCTCGCTCTGAATTTTGTCTTGGAGCCGAGCCGCGCGCGCTTGTGTC 1259
Qy 1557 TGCTGACGATCGTTCTGTGTGTCTGTCTGCTGACTGTGTCTGTCTGTCTGTCTGTCTGAAAT 1616
Db 1260 TGCTGACGATCGTTCTGTGTGTCTGTCTGCTGACTGTGTCTGTCTGTCTGTCTGTCTGAAAT 1319
Qy 1617 ATGGGCCCGGGCTAGACTGTTTACCCTCCCTTAAAGTTTGAACCTTAGGCTCACTGGAAAGAT 1676
Db 1320 ATGGGCCCGGGCTAGACTGTTTACCCTCCCTTAAAGTTTGAACCTTAGGCTCACTGGAAAGAT 1379
Qy 1677 GTCAGCGGATCGCTCAACAGCTCGGTAGATGTCAAGAGAGAGAGTGTGGTTTACCTTC 1736
Db 1380 GTCAGCGGATCGCTCAACAGCTCGGTAGATGTCAAGAGAGAGAGTGTGGTTTACCTTC 1439
Qy 1737 TGCTCTCAGAAATGGCCCAACTTTAAAGTGGCGGAGACGCGACCTTTAAACCGA 1796
Db 1440 TGCTCTCAGAAATGGCCCAACTTTAAAGTGGCGGAGACGCGACCTTTAAACCGA 1499

```


Db 3554 AGCTTAAGTAAAGCCATTTTCAAGGCATGGAATAATACATTAACATGAGATAGAGAAGTT 3613
 Qy 4012 CAGATCAAGGTTAGAAACAGA - GAGACAGCAGAGATATATGGCCAAACAGGATATCTGTGGT 4070
 Db 3614 CAGATCAAGGTCAGGAAACAGATGGAACAGCTGAAATATGGCCAAACAGGATATCTGTGGT 3673
 Qy 4071 AAGCAGTTCTGCCCC - GCTCAGGGCCAGAACAGTGTGGAACAGAGAAATATGGCCAAA 4129
 Db 3674 AAGCAGTTCTGCCCCGCTCAGGGCCAGAACAGATGGAACAGCTGAAATATGGCCAAA 3733
 Qy 4130 CAGGATATCTGTGGTAAGCAGTTCTGCCCCGCTCAGGGCCAGAACAGATGTTGCCCA 4189
 Db 3734 CAGGATATCTGTGGTAAGCAGTTCTGCCCCGCTCAGGGCCAGAACAGATGTTGCCCA 3793
 Qy 4190 GATCGGTCCTCGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 4249
 Db 3794 GATCGGTCCTCGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCTCAA 3853
 Qy 4250 GGACCTGAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGGTTCTCGCTTCTG 4309
 Db 3854 GGACCTGAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTCGGTTCTCGCTTCTG 3913
 Qy 4310 TTGGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGC 4369
 Db 3914 TTGGCGGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGC 3973
 Qy 4370 CAGTCTCCGATPAGACTCGTCGCCCGGGTACCCTGTCTCAATAAACCTCTTGCAGT 4429
 Db 3974 CAGTCTCCGATPAGACTCGTCGCCCGGGTACCCTGTCTCAATAAACCTCTTGCAGT 4033
 Qy 4430 TGCATCCGATCGTGGTCTCGCTGTCTTGGAGGGTCT - CTCGTAGTGAATGACTACC 4488
 Db 4034 TGCATCCGATCGTGGTCTCGCTGTCTTGGAGGGTCTCTCTGAGTGATGACTACC 4093
 Qy 4489 CGTCAGCGGGTCTTTCAGTTTCTCCCACTACACAGTCTCACTAACATTCCTGATGTG 4548
 Db 4094 CGTCAGCGGGTCTTTCAGTTTCTCCCACTACACAGTCTCACTAACATTCCTGATGTG 4113
 Qy 4549 CCGCAGGACTCCGTCAGCCCGGTTTGTGTTTATAAATAAAGCAAGTGTCCCT 4608
 Db 4114 ----- 4113
 Qy 4609 TCAAGCCAGACTACATCTGACTCTCGCTTTATAAAGAATGTGGAAGGCTCTGTGGA 4668
 Db 4114 ----- 4113
 Qy 4669 CTATCGCCACACGACTTTTAAAGATTTTATGCTCTCGATGAGGATTTAGTCAATC 4728
 Db 4114 ----- 4113
 Qy 4729 TATCCTCGTCTATTTTGTGGCTTCTCGGTATTTTAAATTTCTAGTTTTCACCTCCCTTCC 4788
 Db 4114 ----- 4113
 Qy 4789 TGAGAGCAGCGGATTCAGAGTAGTTAATACTCTGAGGGCAGGCTTCTGTGAAAGGTT 4848
 Db 4114 -----TCCGACTTGTGTCTCGCTCTGGGA 4141
 Qy 4849 GCCTGGGCTCAGTGTGAGATTTGGCAATAAAGGGTCTGCCCCCTGTGTACAGACAGA 4908
 Db 4142 GGGTCTCTCTGAGTGATGTACTACCCCTCAGCGGGGTCTTCACATG - 4189
 Qy 4909 TCGGAATCTAGATGTCATCTCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCAT 4968
 Db 4190 ----- 4189
 Qy 4969 CTTTGCTTAGAGATCCTACCGCGAGCATCGTGGCGGGTACCGAGCTCGAATTCGT 5028
 Db 4190 ----- 4189
 Qy 5029 AATCATGTCTAGTGTCTGTTGTAATTTGTTATCCGCTCACAATTCACACAACA 5088
 Db 4190 -----CAGCATGTATCAAAATTAAATTTGGTTTTTTTCTTAAGTATTACATTA 4238

Qy 5089 TACGAGCGGAAAGCATAAAGTGTAAAGCCTCGGGTGCCTAATGAGTGAAGTAACTCACAT 5148
 Db 4239 AA----- 4240
 Qy 5149 TAATTTGGTTCGGCTCACTGCCCGCTTTCAGTTCGGGAAACCTGTCGTGCGCAGCTGCATT 5208
 Db 4241 -----TGGCCATAGTTGCATT 4256
 Qy 5209 AATGAATCGGCCAAACGGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTCCGCTTCT 5268
 Db 4257 AATGAATCGGCCAAACGGCGGGGAGAGCGGTTTTCGTTATTTGGCGCTCTTCCGCTTCT 4315
 Qy 5269 CGCTCACTGACTCGCTCGCTCGCTTCGGTTCGGCTTCGGCGAGCGGTATCAGCTCACTCAA 5328
 Db 4316 CGCTCACTGACTCGCTCGCTCGCTTCGGTTCGGCTTCGGCGAGCGGTATCAGCTCACTCAA 4375
 Qy 5329 AGGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGAGGAAAGAAACATGTGAGCAA 5388
 Db 4376 AGGCGGTAATACGGTTATCCACAGANTCAGGGGATACCGAGGAAAGAAACATGTGAGCAA 4435
 Qy 5389 AAGGCCAGAAAAGGCCAGGAAACCGTAAAGAGCGGTTTCTGTCGCGTTTTTCCATAGGC 5448
 Db 4436 AAGGCCAGAAAAGGCCAGGAAACCGTAAAGAGCGGTTTCTGTCGCGTTTTTCCATAGGC 4495
 Qy 5449 TCCGCCCTCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGA 5508
 Db 4496 TCCGCCCTCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGA 4555
 Qy 5509 CAGGACTATTAAGATACGAGCGTTTCCCTGCGGAGTCCCTCGTGGCTCTCTCTGTTTC 5568
 Db 4556 CAGGACTATTAAGATACGAGCGTTTCCCTGCGGAGTCCCTCGTGGCTCTCTCTGTTTC 4615
 Qy 5569 CGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAGAGCGTGGCGTTT 5628
 Db 4616 CGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGAGAGCGTGGCGTTT 4675
 Qy 5629 CTCACTAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCGCTCCAGGTGGGCT 5688
 Db 4676 CTCACTAGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTCTGCTCGCTCCAGGTGGGCT 4735
 Qy 5689 GTGTGACGAAACCCCGCTTCAGCCGACCGCTCGGCTTATTCGGTAACTATCTGCTCTTG 5748
 Db 4736 GTGTGACGAAACCCCGCTTCAGCCGACCGCTCGGCTTATTCGGTAACTATCTGCTCTTG 4795
 Qy 5749 AGTCCAAACCGGTAAAGACAGACTTATCGCACCTGCGAGCAGCCACTGGTAAACAGGATTA 5808
 Db 4796 AGTCCAAACCGGTAAAGACAGACTTATCGCACCTGCGAGCAGCCACTGGTAAACAGGATTA 4855
 Qy 5809 GCAGAGCGAGGTATGTAGGGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCT 5868
 Db 4856 GCAGAGCGAGGTATGTAGGGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCT 4915
 Qy 5869 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGGAATA 5928
 Db 4916 ACCTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAAGCCAGTTACCTTCGGAATA 4975
 Qy 5929 GAGTTGTAGCTCTTGTATCCGGCAACAAACACCGCTGTGAGCGGTGGTTTTTGT 5988
 Db 4976 GAGTTGTAGCTCTTGTATCCGGCAACAAACACCGCTGTGAGCGGTGGTTTTTGT 5035
 Qy 5989 GCAAGCAGCAGATTTACCGCGCAAAAAAGAGTCTCAAGAGATCTCTTGTATCTTTTCTA 6048
 Db 5036 GCAAGCAGCAGATTTACCGCGCAAAAAAGAGTCTCAAGAGATCTCTTGTATCTTTTCTA 5095
 Qy 6049 CGGGGTCTGACGCTCAGTGGAAACGAAACTCACTTAAGGGATTTGGTCAATGATPAT 6108
 Db 5096 CGGGGTCTGACGCTCAGTGGAAACGAAACTCACGTTAAGGGATTTGGTCAATGATPAT 5155
 Qy 6109 CAAAAAGGATCTTCCACCTAGATCTTTTAAATTTAAATGAAGTTT - TAAATCAATCT 6165
 Db 5156 CAAAAAGGATCTTCCACCTAGATCTTTTAAATTTAAATGAAGTTTGGCAATCAATCT 5215

QY 1677 GTCGAGCGGATCGCTCAAAACAGTCGGTAGATGTCAAGAGAGACGTTGGTTACCTTC 1736
Db |||||
QY 1737 TGCTCTGAGAAATGSCCAACCTTTAACTCGGATGGCGCGAGACGGCACCTTTAAACGA 1796
Db |||||
QY 1797 GACCTCATCACCCAGGTTAAGATCAAGGTCTTTTCACTGGCCCGCATGGACACCCAGAC 1856
Db |||||
QY 1857 CAGTCCCTTACATCGTGACCTGGGAAGCTTTGGCTTTTGAACCCCTTCCCTGGGTCAAG 1916
Db |||||
QY 1917 CCCCTTGTACACCTTAAGCCTCGCCTCTCTCTCCATCGGCCCGCTCTCCCTTT 1976
Db |||||
QY 1977 GAACCTCTCGTTCCAGCCCGCTCGATCTCCCTTTATCCAGCCCTCACTCTCTCTA 2036
Db |||||
QY 2037 GCGCCCCCATATGGCCATATGAGATCTTATATGGGCAACCCCGCCCTTTGTAACCTTC 2096
Db |||||
QY 2097 CTTGACCTTGATGACAGAGATTACTTAAAGCCCTCTCTCCAGCTCACTTACAGGCT 2156
Db |||||
QY 2157 -TCTACTTAGTCCAGACAGAGTCTGGAGACTCTGGCGGAGCTTCAAGAACAACTG 2215
Db |||||
QY 2216 GACCGACCGGTGTTACCTCACCTTTACCGAGTGGCGGACACAGTGTGGGTCCGCGGAC 2275
Db |||||
QY 2276 CAGACTAAGAACCTTAGAACCTCGCTGGAAAGGACCTTACAGTCTGTGACACCCCTC 2335
Db |||||
QY 2336 ACCGCCCTCAAAGTAGAGCGCATCGCAGCTTGGATACACGCCGCCAGCTGAAGGCTGCC 2395
Db |||||
QY 2400 ACCGCCCTCAAAGTAGAGCGCATCGC -GCTTGGATACAGCGCCCGCAGTGAAGGCTGCC 2098
Db |||||
QY 2396 GACCCCGGGGTGGACCACTCTTAGACTGGGATGGAGCTGTATCATCTCTTCT 2455
Db |||||
QY 2456 TGGTAGMAACAGCTACAGGTGTCACTCCGACATCCAGCTGACCCAGAGCCAAAGCAGCC 2515
Db |||||
QY 2516 TGAGCGCAGCGTGGGTGACAGGTGACCATCCTGTAAAGCCAGTCAGGATGTGGTA 2575
Db |||||
QY 2576 CTTCTGTAGCTTGGTACCAGCAGAGCGCAGGTAAAGGCTCCAAAGTGTGTATCTACTGGA 2635
Db |||||
QY 2636 CATCAACCCGCAACATGTTGTGCAAGAGATTCAAGCGGTAGCGGTAGCGGTACCGACT 2695
Db |||||
QY 2696 TCACCTTACCATCAGCAGCTCCAGCCAGAGGACATGCCACCTACTACTTGCAGCAAT 2755
Db |||||
QY 2755 GCA GTTCTCTGGAAGCTTCTTGAAGACAAACAGCTGTGTAGCGACCTT----- 2408

QY 2756 ATAGCCTTATCGTGTCTGGCCCAAGGGACCAAGGTGGAAATCAACGAGGTGGCTCAG 2815
Db |||||
QY 2816 GATCGGTGGATCGGCTCTGGTCTCAGATCGGAGGTCCAACTGGTGGAGAGCGGTG 2875
Db |||||
QY 2875 GAGGTGTTGTGCAACCTTGGCCGCTCTTGGCCCTGTCTGTCTCGGCATCTGGCTTCGATT 2935
Db |||||
QY 2935 TGAGTTGGATAGTTGTGGAAGAGTCAATGGCTCTCTCTCAAGCGTATTCAACAGGGGC 2971
Db |||||
QY 2971 TCACCACATATTTGATGAGTTGGGTGAGACAGGCACTCTGGAAGAGTTTGGAGTGAATG 2995
Db |||||
QY 2995 TGAAGGATGCCAGAGGTACCCCATTTGTATGGGATCTGATCTGGGGCTCGGTGCACAT 2631
Db |||||
QY 2631 GAGAAATTCATCCAGATAGCAGTACGATTAACTATGCGCGTCTCTTAAGGATAGATTTA 3055
Db |||||
QY 3055 GCTTTACATGTGTTTGTGAGGTTTAAACACGCTAGGCCCCCGAACACACGGGACGT 2691
Db |||||
QY 2691 CAATATCGGAGACAAAGCAACATTTGTTCTCTGMAATGGACAGCCTGAGACCG 3115
Db |||||
QY 3115 GGTTTTCTTTGAAAAACAGATGATATATG-----GGGATCCACCG 2736
Db |||||
QY 2736 AAGACACGGGTCTATTTTGTCAAGCCTTTACTTGGCTTCCCTGGTTTGTCTTAT 3175
Db |||||
QY 3175 TCGCACCATGTTGAGCAAGGGCGAGGAGCTGTTTACCGGGTG -GTGCCCATCTCGT 2794
Db |||||
QY 2794 GGGGCCAAGGACCCCGGTACCGTCTCCAGTGTCTAAGCCACCAAGCCAGCGGCGC 3235
Db |||||
QY 3235 CGAGCTGGAGCGGACGTAACCGGCCAAAGTTTCCAGCTGTCCGGCGAGGGCGGCA 2854
Db |||||
QY 2854 GACCACCAACACCGCGCCCAACATCGCTCGCAGCCCTGTCTCTCGCCAGAGGGCG 3295
Db |||||
QY 3295 TGCCACCTACGGCAA -GCTGACCTGAGTTTCTGCAACACCGGAGTGTGCCGTGC 2913
Db |||||
QY 2913 CTGGCCCACTCTGTGACCAACCTGACCTAGCGGTGAGTGTCTTACCGCTTACCCCG 2973
Db |||||
QY 2973 AACTCTCTACTCTGTGATGGAATCTCTTCTATGTTGTGTCTATCTACTGCTTGT 3415
Db |||||
QY 3415 ACCACATGAAGCAGCAGN-----CTTCTTCAAGTCCGCCATGCCGAAGGCTACG 3024
Db |||||
QY 3024 TCTCTGAGAGTGAAGTTTCCAGGAGCGCAGAGCCCGCGTACAGCAGGCGCAGAAC 3475
Db |||||
QY 3475 TCCAGGAGCGCAACATTTCTTCAAGGACGAGCGCAACTACAGACCCGCCCGAGGTGA 3084
Db |||||
QY 3084 AGCTCTATAACGAGCTCAATCTTAGGACGAAAGAGGAGTACGATGTTTGGACAGAGAC 3535
Db |||||
QY 3535 AGTTCCGAGGGCGACACCTCTGTGAACCGCATCGAGCTGAAGGCATCGACTTCAAGGAG 3144
Db |||||
QY 3144 GTGGCCGGGACCTTGAGATGGGGGAAAGCGGAGAGAGAACCTCTCAGG---AAGGCC 3592
Db |||||
QY 3592 ACGGCAACATCTCTGGGCGCAAGCTGGAGTACAACTACAAACAGCCACAACTCTATATCA 3204
Db |||||
QY 3204 TGTACATGATGACGAAAGATAAGATGCGGAGGCTTACAGTGAGATTGGGATGAAG 3652
Db |||||
QY 3652 TGGCCGCAAGAGAGAACCGGCATCAAGGTGAATCTTCAAGTCCGCCCAACATCTCGAG 3264
Db |||||
QY 3264 GCGAGCCCGGAGGGCAAGGGCGACGATGGCTTTTACAGGGTCTCAGTACAGCCACCA 3712
Db |||||
QY 3712 ACGCAGCGTGACAGTCCGCGACCACTTACAGCAGAACACCCCATCGGACAGCCCGCG 3324
Db |||||
QY 3324 AGGACACTTACGAGCCCTTCAATGACGGCTTGGCCCTCTCGCTAACTCGACGGGGCG 3772
Db |||||
QY 3772 TGCTGTCTGCGCAACCACTTACCTGAGCACCCAGTCCGCCCTGAGCAAGAGACCCCAACG 3384
Db |||||
QY 3384 CGGATCCGGATAGTCCAAATTTGTTAAAGACAGATATCAGTGTCTCCAGGCTCTAGTTTT 3832
Db |||||
QY 3832 AGAAGCGCGATCA-----CATGGTCTCTGTGAGTTTGTGACCGCCCGCGGAT 3433
Db |||||
QY 3433 GACTCAACAATATCACCGACTGAAGCCTATAGAGTACGAGCCATAGATAAATAAAGAT 3892

3434 CACTCTCGGCATGGACGAGCTGTACAAAGTAAAGCGCGCTCGACGATAAAATAAAGAT 3493
 3893 TTTATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTGTGGCAAGCT 3952
 3494 TTTATTTAGTCTCCAGAAAAGGGGGAATGAAAGACCCACCTGTAGTGTGGCAAGCT 3553
 3953 AGCTTAAGTAAAGCCCAATTTTGAAGGCATGG-AAAATACATAAATCAGAGATAGAGAGTT 4011
 3554 AGCTTAAGTAAAGCCCAATTTTGAAGGCATGGAAAATAATACATAAATCAGAGATAGAGAGTT 3613
 4012 CAGATCAAGGTTAGGAACAGA-GAGACAGAGAAATATGGGCCAAACAGAGATATCTGTGT 4070
 3614 CAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGAGATATCTGTGT 3673
 4071 AAGCAGTTCCTGCCCC-GCTCAGGGCCAAAGACAGTTGGAAACAGGAGAAATATGGGCCAAA 4129
 3674 AAGCAGTTCCTGCCCCCGCTCAGGGCCAAAGACAGATGGAACAGCTGAATATGGGCCAAA 3733
 4130 CAGGATATCTGTGGTAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGACAGATGTTGCCCA 4189
 3734 CAGGATATCTGTGGTAAGCAGTTCTGCCCCCGCTCAGGGCCAAAGACAGATGTTGCCCA 3793
 4190 GATGCGGTCCGCGCTCAGAGTTTCTAGAGAACATCAGATGTTTTCAGGGTGGCCCCAA 4249
 3794 GATGCGGTCCGCGCTCAGAGTTTCTAGAGAACATCAGATGTTTTCAGGGTGGCCCCAA 3853
 4250 GGACCTGAATGACCTGTGCTTATTTGAATACCAATCAGTTCGCTTCTCGCTTCG 4309
 3854 GGACCTGAATGACCTGTGCTTATTTGAATACCAATCAGTTCGCTTCTCGCTTCG 3913
 4310 TTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGCG 4369
 3914 TTGCGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACACCCCTCACTCGGCGCG 3973
 4370 CAGTCTCCGATAGACTGCGTCCCGGTAACCGGTGTTCTCAATAAACCCCTTTGCAAT 4429
 3974 CAGTCTCCGATAGACTGCGTCCCGGTAACCGGTGTTCTCAATAAACCCCTTTGCAAT 4033
 4430 TGCATCCGACTGCTGCTGCTGCTTCTTGGAGGGTCT-CTCTGAGTATGACTACC 4488
 4034 TGCATCCGACTGCTGCTGCTGCTTCTTGGAGGGTCTCTCTGAGTATGACTACC 4093
 4489 CGTACGCGGGTCTTTCAGTTCCTCCACCTACACAGGTCTCACTAAACATTCCTGATGTG 4548
 4094 CGTCCGCGGGCTTTTCAT- 4113
 4549 CCGCAGGGAACCTCGTCAGCCCGGTTTTTGTATATAAATAATGCAAGAACAGTGTTCCT 4608
 4114 - 4113
 4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTATATAAAGATGTTGAAGGGCTCTGTGA 4668
 4114 - 4113
 4669 CTATCTGCCACACGACTTTTAAAGATTTTATAGCTTCTGCTGATGAGGATTTAGTCAATC 4728
 4114 - 4113
 4729 TATCCTCGTATATTTTGTGGCTTCTCGTATTTTAAATTTCTAGTTTTCACCTCCCTCC 4788
 4114 - 4113
 4789 TGAGAGCAGCGGATTCAGAGTAGTTAATACCTTGAGGGCAGGCTTCTGTGAAAAGTTT 4848
 4114 - 4113
 4849 GCTCGGCTCAGTGTGAGATTTTGCCATAAAGGGGCTCCCTGCTGTGTACAGACAGA 4908
 4142 GGGTCTCCCTCTGAGTGTGACTACCCGTCAGGGGGGTCTTCAATG- 4189
 4909 TCGGAATCTAGAGTGATCTACAGAGTCCCGCGGTTCCGGGGCTCTGTATCTCAGGGCAT 4968

4190 - 4189
 4969 CTTTGCTTAGAGATCCTCTACGCGGACGCACTCGTGGCGGGTACCGAGCTCGAATTCCT 5028
 4190 - 4189
 5029 AATCATGGTTCATAGCTGTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCACACAA 5088
 4190 - 4189
 5089 TAGAGCCGGAAGCATAAAGTGTAAAGCTTGGGGTCCCTAATGAGTGAAGTAACTACAT 5148
 4239 AA- 4240
 5149 TAAATGCGTGTGGCTCACTGCGCGCTTTTCCAGTCGGGAAACCTGCTGCGCAGCTGCAT 5208
 4241 - 4256
 5209 AATGAATCGGCCAAACGCGGGGAGAGCGGTTTTCGTAATTTGGCGCTCTTCCGCTTCCT 5268
 4257 AATGAATCGGCCAAACGCGGGGAGAGCGGTTTTCGTAATTTGGCGCTCTTCCGCTTCCT 4315
 5269 CGCTCACTGACTCGCTGCGCTCGGCTCGGCTCGGCGAGCGGTATCAGCTCACTCAA 5328
 4316 CGCTCACTGACTCGCTGCGCTCGGCTCGGCTCGGCGAGCGGTATCAGCTCACTCAA 4375
 5329 AGGCGGTAAACGCTTATCCACAGAAATCAGGGGATACGCGGAAAGAACATGTGAGCAA 5388
 4376 AGGCGGTAAACGCTTATCCACAGAAATCAGGGGATACGCGGAAAGAACATGTGAGCAA 4435
 5389 AAGGCGGTAAACGCTTATCCACAGAAATCAGGGGATACGCGGAAAGAACATGTGAGCAA 5448
 4436 AAGGCGGTAAACGCTTATCCACAGAAATCAGGGGATACGCGGAAAGAACATGTGAGCAA 4495
 5449 TCCGCGCGCTGAGCGAGCATCAAAAAATCAGAGCTCAAGTCAAGTGGGGAACCCGA 5508
 4496 TCCGCGCGCTGAGCGAGCATCAAAAAATCAGAGCTCAAGTCAAGTGGGGAACCCGA 4555
 5509 CAGGACTATAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTTTC 5568
 4556 CAGGACTATAAGATACAGAGGTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTTTC 4615
 5569 CGACCTCGCGCTTACCGGATACCTGCTCGGCTTCTCCCTTCGGGAAAGCGTGGCGCTTT 5628
 4616 CGACCTCGCGCTTACCGGATACCTGCTCGGCTTCTCCCTTCGGGAAAGCGTGGCGCTTT 4675
 5629 CTATAGCTCAAGCTGTAGGTATCTCAGTTCGCTGAGTTCGCTTCAAGTGTGGCT 5688
 4676 CTATAGCTCAAGCTGTAGGTATCTCAGTTCGCTGAGTTCGCTTCAAGTGTGGCT 4735
 5689 GTGTGACGAAACCCCGCTCAGCCCGCTGAGCTGCGCTTATCCGCTAATCTCTGTTG 5748
 4736 GTGTGACGAAACCCCGCTCAGCCCGCTGAGCTGCGCTTATCCGCTAATCTCTGTTG 4795
 5749 AGTCCAAACCCGCTAAGACACGACTTATCCGCTGAGCTGAGCGGCTGTTAAGAGATTA 5808
 4796 AGTCCAAACCCGCTAAGACACGACTTATCCGCTGAGCTGAGCGGCTGTTAAGAGATTA 4855
 5809 GCAGAGCAGGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCT 5868
 4856 GCAGAGCAGGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACTACGGCT 4915
 5869 ACACCTAGAGACACGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTCCTCGGAAAA 5928
 4916 ACACCTAGAGACACGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTCCTCGGAAAA 4975
 5929 GAGTGTGAGTCTTGTATCCGCAAAACAAACCCGCTGAGCGGCTGTTTGTGTTT 5988
 4976 GAGTGTGAGTCTTGTATCCGCAAAACAAACCCGCTGAGCGGCTGTTTGTGTTT 5035
 5989 GCAAGCAGCAGATTAACGCGCAAAAGAGATCTCAAGAGATCTCTTGTGATCTTTTCTA 6048
 5036 GCAAGCAGCAGATTAACGCGCAAAAGAGATCTCAAGAGATCTCTTGTGATCTTTTCTA 5095


```
QY 6049 CCGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCTAGATTTAT 6108
Db 5096 CCGGCTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAGGATTTTGGTCTAGATTTAT 5155
QY 6109 CAAAAGGATCTTACACCTAGATCTTTTAAATTAATAATGAAGTTT---TAAATCAATCT 6165
Db 5156 CAAAAGGATCTTACACCTAGATCTTTTAAATTAATAATGAAGTTTGGCAATCAATCT 5215
QY 6166 AAAGTATATAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACCTA 6225
Db 5216 AAAGTATATAGTAACCTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACCTA 5275
QY 6226 TCTAGCGATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAA 6285
Db 5276 TCTAGCGATCTGCTATTTGCTTCATCCATAGTTGCTGACTCCCGCTGCTGTAGATAA 5335
QY 6286 CTAGATACGGAGGGCTTACCATCTGCGCCAGTCTGCAATGCTATATACCGGAGCCAC 6345
Db 5336 CTAGATACGGAGGGCTTACCATCTGCGCCAGTCTGCAATGCTATATACCGGAGCCAC 5395
QY 6346 GCTCACCCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAAGGCCGAGCGAGAA 6405
Db 5396 GCTCACCCTCCAGATTTATCAGCAATAAACCCAGCCAGCGGAAGGCCGAGCGAGAA 5455
QY 6406 GTGCTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTCGGGAAGCTAGAG 6465
Db 5456 GTGCTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTCGGGAAGCTAGAG 5515
QY 6466 TAAGTAGTCCCGAGTTAATAGTTTGGCAAGTTTGGCAATGCTGCTACAGGC-TCGTGG 6524
Db 5516 TAAGTAGTCCCGAGTTAATAGTTTGGCAAGTTTGGCAATGCTGCTACAGGCATCGTGG 5575
QY 6525 TGTCAAGCTCTGCTGTTGTTGTTGCTTCAATTCAGCTCCGTTTCCCAAGTCAAGGCGAG 6584
Db 5576 TGTCAAGCTCTGCTGTTGTTGTTGCTTCAATTCAGCTCCGTTTCCCAAGTCAAGGCGAG 5635
QY 6585 TTACATGATCCCACTGTTGTGCAAAAAGCGTTAGTCTCTCGTCTCCGATCGTTG 6644
Db 5636 TTACATGATCCCACTGTTGTGCAAAAAGCGTTAGTCTCTCGTCTCCGATCGTTG 5695
QY 6645 TCAGAAGTAAAGTTGGCCAGTGTATACATCATGTTTATGGCAGCATGATTAATCTC 6704
Db 5696 TCAGAAGTAAAGTTGGCCAGTGTATACATCATGTTTATGGCAGCATGATTAATCTC 5755
QY 6705 TTACTGTCTATGCTCCGTAAGATGCTTTCTGTGACTGGTGAAGTCTCAACCAAGTAT 6764
Db 5756 TTACTGTCTATGCTCCGTAAGATGCTTTCTGTGACTGGTGAAGTCTCAACCAAGTAT 5815
QY 6765 TCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAATA 6824
Db 5816 TCTGAGATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATAATA 5875
QY 6825 CCGGCCCATAGACGAACTTTAAAGTGTCTATCATATGGAATAAGTTCTTCCGGGCGAA 6884
Db 5876 CCGGCCCATAGACGAACTTTAAAGTGTCTATCATATGGAATAAGTTCTTCCGGGCGAA 5935
QY 6885 AACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGAGTAAACCACTGCTGACCCCA 6944
Db 5936 AACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGAGTAAACCACTGCTGACCCCA 5995
QY 6945 ACTGATCTTACGATCTTTTAACTTTTACACAGCGTTTCTGGGTGAGCAAAAACAGAGGC 7004
Db 5996 ACTGATCTTACGATCTTTTAACTTTTACACAGCGTTTCTGGGTGAGCAAAAACAGAGGC 6055
QY 7005 AAAATGCCGCAAAAAGGAAATAAGGCGCACACGGAATGTTGAATACTCATCTCTTCC 7064
Db 6056 AAAATGCCGCAAAAAGGAAATAAGGCGCACACGGAATGTTGAATACTCATCTCTTCC 6115
QY 7065 TTTTCAATATATTAAGACATTTATCAGGTTTATGTTCTCATGAGCGGATACATATTTG 7124
Db 6116 TTTTCAATATATTAAGACATTTATCA-GGTTATGTTCTCATGAGCGGATACATATTTG 6174
```

```
QY 7125 AATGTATTTAGAAAAATAAAACAATAGGGTTTCGGCGCACATTTTC 7169
Db 6175 AATGTATTTAGAAAAATAAAACAATAGGGTTTCGGCGCACATTTTC 6219
```

RESULT 13

```
US-09-963-247A-4
; Sequence 4, Application US/09963247A
; Publication No. US20040214155A1
; GENERAL INFORMATION:
; APPLICANT: Ferrick, David A.
; APPLICANT: Swift, Susan E.
; APPLICANT: Armstrong, Randall
; APPLICANT: Fox, Bryan
; TITLE OF INVENTION: Methods and Compositions for Screening for Modulators and Ige Syn
; FILE OF INVENTION: Secretion and Switch Rearrangement
; FILE REFERENCE: A-66038-1/RMS/JJD/DLR
; CURRENT APPLICATION NUMBER: US/09/963,247A
; CURRENT FILING DATE: 2002-05-07
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 6219
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-963-247A-4
```

```
Query Match 40.3%; Score 3082.2; DB 11; Length 6219;
Best Local Similarity 73.7%; Pred. No. 0;
Matches 4587; Conservative 0; Mismatches 958; Indels 680; Gaps 25;
```

```
QY 958 CCGAGCTCAATAAAGAGCCCAACCCCTCAGTCGGGGCCAGTCTCCGATAGACTG 1017
Db 662 CAGAGCTCAATAAAGAGCCCAACCCCTCAGTCGGGGCCAGTCTCCGATAGACTG 721
QY 1018 CGTCGCCGGGTACCCGCTATTCCCAATAAAGCCCTCTTGTCTTTGCATCCGAATCGTGA 1077
Db 722 AGTCGCCGGGTACCCGCTATTCCCAATAAAGCCCTCTTGTGAG-TTGCATCCGACTTGTGT 780
QY 1078 CTCGCTGATCTTGGGAGGGTCTCCTCAGATTTGATTTGACTGCCCC-ACCTCGGGGGTCTTT 1136
Db 781 CTGCTGTTCTTGGAGGGTCTCCTCTGAGTATTGACTACCCGTGAGCGGGGTCTTT 840
QY 1137 CATTTGGAGGTTCCACCGAGATTTTGAGACCCCTGCCCCAGGACACCGACCCCGGCC 1196
Db 841 CATTTGGGGGCTCGTCGGGATCGGAGACCCCTGCCCCAGGACCCCGA-CCCAACCACC 899
QY 1197 GGGAGGTAAGTCGCCAGCACTTATCTGTCTGTCCGATGTTGCTAGTGTCTATGACTG 1256
Db 900 GGGAGGTAAGTCGCCAGCACTTATCTGTCTGTCCGATTTGTCTCGATTTGTCTATGACTG 959
QY 1257 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCCGTG 1316
Db 960 ATTTTATGCGCTCGCTCGGTACTAGTTAGTAACTAGTCTGTATCTGGCGGACCCCGTG 1019
QY 1317 GTGAACTGACGATTCGGAAACACCCCGCGCGAAACCTTGGAGAGCTGCCAGGACTTCG 1376
Db 1020 GTGAACTGACGATTCGGAAACACCCCGCGCGAAACCTTGGAGAGCTGCCAGGACTTCG 1079
QY 1377 GGGCGCTTTTGTGGCCCGACCTGAGTCTCTAAATCCGATCCGATTTAGGACTCTTTGGT 1436
Db 1080 GGGCGCTTTTGTGGCCCGACCTGAGTCTCTAAATCCGATCCGATTTTGGACTCTTTGGT 1139
QY 1437 GCACCCCTTTAGAGGAGGATATGTGTTCTGTGTAGGACGAGAACTAAAAACAGTTC 1496
Db 1140 GCACCCCTTTAGAGGAGGATATGTGTTCTGTGTAGGACGAGAACTAAAAACAGTTC 1199
QY 1497 CCGCTCCGTCTGAATTTTTCGTTTGGGACCCGAACCGCGCGCGCTCTTCTC 1556
```

Db 1200 CGGCTCCGCTGTAATTTTGGTTTGGTTCGTTTGGGACCGGAGCGCGCGGCTCTTGTC 1259
 Qy 1557 TGCTGAGCATCGTTCTGTGTGTCTCTCTCTGACTGTGTCTCTGTATTTGTCTGAAT 1616
 Db 1260 TGCTGAGCATCGTTCTGTGTGTCTCTCTCTGACTGTGTCTCTGTATTTGTCTGAAT 1319
 Qy 1617 ATGGGCGGGGTAGACTGTATACCACTCCCTTAAGTTTACCTTAGGTCACTTGGAAAGAT 1676
 Db 1320 ATCGGCGGGGAGACTGTATACCACTCCCTTAAGTTTGAAGTTTGAAGTTTGAAGAT 1379
 Qy 1677 GTGAGCGGATCGCTCACACAGTCGAGTAGTGTCAAGAGAGAGCTTTGGTTACCTTC 1736
 Db 1380 GTGAGCGGATCGCTCACACAGTCGAGTAGTGTCAAGAGAGAGCTTTGGTTACCTTC 1439
 Qy 1737 TGCTCTGAGATGCGCAACCTTTAAGTCGAGTGGCCGCGAGACGGGACCTTTAACCGA 1796
 Db 1440 TGCTCTGAGATGCGCAACCTTTAAGTCGAGTGGCCGCGAGACGGGACCTTTAACCGA 1499
 Qy 1797 GACCTCATCACAGGTTAAGATCAAGTCTTTTCACTGGCCCGCATGGACACCCAGAC 1856
 Db 1500 GACCTCATCACAGGTTAAGATCAAGTCTTTTCACTGGCCCGCATGGACACCCAGAC 1559
 Qy 1857 CAGGTCCCTACATCGTACCTGGGAAGCTTTGGCTTTTGAACCCCTCCCTGGGTCAAG 1916
 Db 1560 CAGGTCCCTACATCGTACCTGGGAAGCTTTGGCTTTTGAACCCCTCCCTGGGTCAAG 1619
 Qy 1917 CCGTTGTACACCTTAAGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1976
 Db 1620 CCGTTGTACACCTTAAGCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1679
 Qy 1977 GAACCTCTCTCTGACCCCGCTCGATCCCTCTTATCCAGCCCTCACTCTCTCTCTA 2036
 Db 1680 GAACCTCTCTCTGACCCCGCTCGATCCCTCTTATCCAGCCCTCACTCTCTCTCTA 1739
 Qy 2037 GCGGCCCTCATATGGCCATATGAGTCTTATATGGGCAACCCCGCCCTTGTAACTTC 2096
 Db 1740 GCGGCCCTCATATGGCCATATGAGTCTTATATGGGCAACCCCGCCCTTGTAACTTC 1799
 Qy 2097 CTGACCTCTGACATGACAGAGTTACTAAACAGCCCTCTCTCAAGCTCACTTACAGGT 2156
 Db 1800 CTGACCTCTGACATGACAGAGTTACTAAACAGCCCTCTCTCAAGCTCACTTACAGGT 1859
 Qy 2157 -TCTACTTAGTCAGACCAAGTCTGGAGACCTCTGGGCGAGCCTTACCAGAACACTG 2215
 Db 1860 CTCTACTTAGTCAGACCAAGTCTGGAGACCTCTGGGCGAGCCTTACCAGAACACTG 1919
 Qy 2216 GACCGACCGGTGATCTCTCAACCTTACCGAGTCGCGACACAGTGTGGTCCGCGACAC 2275
 Db 1920 GACCGACCGGTGATCTCTCAACCTTACCGAGTCGCGACACAGTGTGGTCCGCGACAC 1979
 Qy 2276 CAGACTAAGAACTTGAACCTCGCTGGAAGAGCCTTACAGTCTGTGTGACCAACCC 2335
 Db 1980 CAGACTAAGAACTTGAACCTCGCTGGAAGAGCCTTACAGTCTGTGTGACCAACCC 2039
 Qy 2335 ACGGCCCTAAGTAGAGCGGATCGAGCTTGGATACACCGCCCGCCAGTGAAGGTGCC 2395
 Db 2040 ACGGCCCTAAGTAGAGCGGATCGG-GCTTGGATACACCGCCCGCCAGTGAAGGTGCC 2098
 Qy 2396 GACCCGGGGTGGACATCTCTAGACTGCCATGGGATGGAGTGTATCATCTCTCTCT 2455
 Db 2099 GACCCGGGGTGGACATCTCTAGACTGCC-----GGATCTGAGGG 2142
 Qy 2456 TGGTAGCAACAGCTACAGGTGCCATCCGACATCCAGTGAACCCAGAGCCCAAGCAGCC 2515
 Db 2143 ATCCACCAACAGGACCCCATTAATTTGAATTCCTGAGCCCGGGGATCCACTAGTT 2202
 Qy 2516 TGAGCGCAGCTGGGTGACAGAGTGAACCATCACCTGTAAAGCCAGTCAAGATGGGTA 2575
 Db 2203 CTAGAGCGAATTAATTCGGTTATTTTCACCA-----TA 2237
 Qy 2576 CTTCTGTAGTTGGTACCAGCAGAGCCAGGTAAAGCTCCAAAGCTGTGATCTACTGGA 2635
 Db 2238 TTGCCGTCTTTTGGCAATGTAGGGCCCGGAACCTGGGCCCTCTCTTTTGAAGCAAT 2297

Qy 2636 CATCCACCGGCGACACTGCTGTGCCAAGAGATTTCAGCGGTAGCGGTACCGACT 2695
 Db 2298 CTTAGGGGTCTTTCCTCTCGCCAAAGAAATGCAAGGTCTGTGAATGTCTGTGAAGNA 2357
 Qy 2696 TCACCTTCACCATCAGCAGCCTCCAGCGAGAGGACATCGCCACCTACTACTGCCACAT 2755
 Db 2358 GCAGTTCTCTGGAAGCTTCTTGAAGACAACAACAGTCTGTAGCGACCTT----- 2408
 Qy 2756 ATAGCTCTATCGGTCTGTGGCCAAAGGACCAAGGTGGAATCAAAAGAGTGGCTCAG 2815
 Db 2409 -----TGCAGGACAGCGAAACCCCACTGCGCCACAGGTCTCTCG 2451
 Qy 2816 GATCGGTGGATCCGCTCTGTGCTCAGGATCGAGGTCCAACTGTGTGAGAGCGGTG 2875
 Db 2452 GCCAAAAGCCAGTGTATAAGATACACCTGCAAGGCGGCAACCCAGTGTCCACGTTG 2511
 Qy 2876 GAGGTGTGTGCAACCTGCGCGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 2935
 Db 2512 TGAGTTGATAGTTGTGGAAGAGTCAATGGCTCTCTCAAGCTATTTCAACAGGGGC 2571
 Qy 2936 TCACCAATATTTGGATAGTTGGGTGAGACAGGACCTGTGGAAGAGTCTTGAAGTATG 2995
 Db 2572 TGAAGGATGCCAGAAAGTACCCCATTTGTATGGGATCTGATCTGGGGCTCTGGTGCACAT 2631
 Qy 2996 GAGAAATTCATCCAGATAGCAGTAACTATGCGCGCTCTCTAAAGATAGATTTA 3055
 Db 2632 GCTTTACATGTGTATGAGGTTTAAAGGCTGTAGGCCCGCCGAAACACCGGAGAGT 2691
 Qy 3056 CAATATCGGAGACAAACCGCAAGAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 3115
 Db 2692 GGTTCCTTTGAAAACACGATGATATG-----GGGATCCACCG 2736
 Qy 3116 AAGACACCGGGTCTATTTTGTGCAAGCTTTACTTGGGTCTCCCTGTGTGTGTATT 3175
 Db 2737 TCGCCACCATGTGTGAGCAAGGCGGAGGAGTGTTCACCGGGTG--GTGCCCATCTGT 2794
 Qy 3176 GGGGCAAGGACCCCGGTCACTGTCTCAGTGTAAAGCCACACGAGCCAGCGCGCG 3235
 Db 2795 CGAGTGGACGGCGAGTAAACGGCCACAAGTTTCAAGCTGTCCGCGAGGGCGAGGCGA 2854
 Qy 3236 GACCAACCAACACCGCGCGCCACCTCGGTCTCGAGCCCTGTCTCTCTCTCTCTCTCT 3295
 Db 2855 TGCCACCTACCGGCA--GCTGACCTTGAAGTTTCACTGTACACACCGGCAAGCTGCCGTGC 2913
 Qy 3296 CTCGCGCAGCGCGGGGCGCAGTGCACAGAGGGGTGGAATTCGCTCTGATTCCTCA 3355
 Db 2914 CTTGGCCCACTCTGTGACCACTCTGACCTTACGGGTGCGAGTCTTCAAGCGCTACCCG 2973
 Qy 3356 AACTCTGTCTCTGTGATGGAATCTCTTCACTATGTTGTCTTCTCTCTCTCTCTCTCT 3415
 Db 2974 ACCACATGAAGCAGCAGCA-----CTTCTTCAAGTTCGCGCCATCGCCGAAGCTAGC 3024
 Qy 3416 TCCTCAGAGTGAAGTTTCAAGAGGCGAGAGCCCGCGCTACAGAGGGCGCAGAAC 3475
 Db 3025 TCCAGAGCGCACATCTTTTCAAGAGAGCGGCAACTAAGACCCCGCGCGAGGTGA 3084
 Qy 3476 AGCTCTATAACGAGCTCAATCTAGGACGAAGAGAGGTAGATGTTTGTGGAAGAGAC 3535
 Db 3085 AGTTTCAGGGCGACACCTCTGTGTAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGG 3144
 Qy 3536 GTGGCGGGGACCTGAGATGGGGGAAAGCGAGAGAGAACCTCTCAGG---AAGGCC 3592
 Db 3145 ACGGCAACATCTCTGGGCGCAAGCTGGAGTACAACTACAAAGCCACACCTCTATATCA 3204
 Qy 3593 TGTAACAATCACTCAGAAAGATAGATGGCGGAGGCTTACAGTGAAGTGGGATGAAAG 3652
 Db 3205 TGCGCGCAAGCAGAAAGACCGCATCAAGGTGAATTTCAAGATCGGCCAACAATCGAGG 3264
 Qy 3653 GCGAGCGCGGAGGGCGCAAGGGGCAAGTGGCGGAGGCTTACAGTGAAGTGGGATGAAAG 3712
 Db 3265 ACGGCGAGCTGCGAGCTCGCCGACCACTACAGCAGAACACCCCATCGCGAGCGGCCCG 3324

QY 3713 AGGACACCTACGAGCCCTTACATGTCAGGCGCTGCCCTCGCTAACTCGACGGCGCG 3772
DB 3325 TGCTGCTGCCGACAACTACTACTGAGCACCAGCTCGCGCTGAGCAAGACCCCAACG 3384
QY 3773 CGGATCCGGATTAGTCCAAATTTGTTAAAGACAGAGGATATCAGTGGTCCAGGCTCTAGTTT 3832
DB 3385 AGAGCGCGATCA-----CATGGTCTGCTGAGTTCGTGACCGCGCGCGGAT 3433
QY 3833 GACTCAACAATATCAACGAGTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGAT 3892
DB 3434 CACTCTCGCATGAGCAGCTGTACAAGTAAAGCGCGCTCGACGATAAATAAAGAT 3493
QY 3893 TTTATTTAGTCTCCAGAAAGAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCT 3952
DB 3494 TTTATTTAGTCTCCAGAAAGAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCT 3553
QY 3953 AGCTTAAGTAAAGCCATTTTGAAGGCTAGG-AAAATACATAACTGAGAAATAGAGAGTT 4011
DB 3554 AGCTTAAGTAAAGCCATTTTGAAGGCTAGG-AAAATACATAACTGAGAAATAGAGAGTT 3613
QY 4012 CAGATCAAGGTTAGGAACAGA-GAGACAGCAGAAATATGGGCCAAACAGGATATCTGTGT 4070
DB 3614 CAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGCCAAACAGGATATCTGTGT 3673
QY 4071 AAGCGTTCTGCCCC-GCTCAGGCCCAAGAACAGTTGGAAACAGAGAAATATGGGCCAAA 4129
DB 3674 AAGCGTTCTGCCCCGCTCAGGCCCAAGAACAGATGGAACAGCTGAATATGGGCCAAA 3733
QY 4130 CAGGATATCTGTGGTAAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCA 4249
DB 3794 GATGGGTCAGCCCTCAGCAGTTCTTAGAGAACCATCAGATGTTTCCAGGGTGCCTCA 3853
QY 4250 GGACTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTGCTTCTGCTTCTG 4309
DB 3854 GGACTGAAATGACCTGTGCTTATTTGAACCTAACCAATCAGTTGCTTCTGCTTCTG 3913
QY 4310 TTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGCGC 4369
DB 3914 TTCCGGCGCTTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCAGTCCGGCGC 3973
QY 4370 CAGTCTCCGATAGACTCGCTCGCCGGGTACCCGCTGTTCTCAATAAACCTCTTGCAGT 4429
DB 3974 CAGTCTCCGATGACTGAGTCCGCGGTACCCGCTGTTCTCAATAAACCTCTTGCAGT 4033
QY 4430 TGCAATCGACTCGTGGTCTCGCTGTTCTTGGGAGGGTCT-CTCTGAGTGAATGACTACC 4488
DB 4034 TGCAATCGACTGTTGGTCTCGCTGTTCTTGGGAGGGTCTCTCTGAGTGAATGACTACC 4093
QY 4489 GTCAGCGGGTCTTTAGTTTCTCCACCTACACAGGCTCTCACTAACATTCCTGATGTG 4548
DB 4094 GTCGCGGGGTCTTTTCAAT----- 4113
QY 4549 CCGCAGGAGCTCCGTCAGCCCGTTTTTGTGTTTATAATAAATGCAAGAACAGTGTCCCT 4608
DB 4114 ----- 4113
QY 4609 TCAAGCCAGACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGGCTCTGTGA 4668
DB 4114 ----- 4113
QY 4669 CTAATCGCCACAGACTTTTAAAGATTTTATGCTCTCGATGAGGAAATTAAGTCAATC 4728
DB 4114 ----- 4113
QY 4729 TATCTCTGCTATTTTGTGGCTTCTCGCTATTTTAAATTTCTAGTTTGCACTCCCTCC 4788
DB 4114 ----- 4113
QY 4789 TCAGAGCAGCGGATTGCAGAGTAGTTAATACTCTGAGGCGAGGCTTCTGTGAAAAAGGTT 4848

DB 4114 -----TCCGACTTGTGTCTCGCTGCCTTGGGA 4141
QY 4849 GCCTGGGCTCAGTGTGAGATTTTGCCATAAAAGGGGCTCTGCCCTCTGTACAGACAGA 4908
DB 4142 GGGTCTCTGTGAGTATTGACTACCGGTACCGGGGTCTTCAATG----- 4189
QY 4909 TCGGAATCTAGAGTGCATCTACAGAGTCCCGCGGTTCGGGGGTCTGATCTCAGGGCAT 4968
DB 4190 ----- 4189
QY 4969 CTTTGCCTAGAGATCCTCTACGCCGAGCATCGTGCCGGGTACCGAGCTCGAATTCGT 5028
DB 4190 ----- 4189
QY 5029 AATCATGTGTATAGTCTGTTTCTGTGAAATTTGTTATCCGCTCACAATTCACAACA 5088
DB 4190 -----CAGCATGTATCAAAATTAATTTGGTTTTTTTCTTAAGTATTACATTA 4238
QY 5089 TACGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTGAAGTAACTCAT 5148
DB 4239 AA----- 4240
QY 5149 TAATTTGGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCTGTCAGCTGCATTT 5208
DB 4241 -----TGGCCATAGTTGCATTT 4256
QY 5209 AATGAATCGGCCAACGCGGGGAGAGCGGTTTTGCTATTTGGGCGCTCTTCCGCTTCT 5268
DB 4257 AATGAATCGGCCAACGCGGGGAGAGCGGTTTTGCTATTTGGGCTCTTCCGCTTCT 4315
QY 5269 CGCTCACTGACTCGCTCGCTCGTTGCGGTCGCGGAGCGGTATCAGCTCACTCAA 5328
DB 4316 CGCTCACTGACTCGCTCGCTCGTTGCGGTCGCGGAGCGGTATCAGCTCACTCAA 4375
QY 5329 AGGCGGTAAATAGGTTATCCAGAAATCAGGGGATACGAGGAAAGAAATGTGAGCAA 5388
DB 4376 AGGCGGTAAATAGGTTATCCAGAAATCAGGGGATAACGAGGAAAGAAATGTGAGCAA 4435
QY 5389 AAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTTCGTGCGTTTTTCCATAGC 5448
DB 4436 AAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGTTCGTGCGTTTTTCCATAGC 4495
QY 5449 TCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGGAAACCCGA 5508
DB 4496 TCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGGAAACCCGA 4555
QY 5509 CAGGACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTCTGTTT 5568
DB 4556 CAGGACTATAAAGATACAGGCGTTTTCCCTGGAAGCTCCCTCGTGGCTCTCTCTGTTT 4615
QY 5569 CGACCTGCGGCTTACCGGATACCTGTCGCGTTTTCTCCCTCGGGAGCGTGGCGTTT 5628
DB 4616 CGACCTGCGGCTTACCGGATACCTGTCGCGTTTTCTCCCTTCGGGAGCGTGGCGTTT 4675
QY 5629 CTCTAGTCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTTCCGCTCCAGCTGGGCT 5688
DB 4676 CTCTAGTCTACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTTCCGCTCCAGCTGGGCT 4735
QY 5689 GTGTGACAGAACCCCGCTTACGCCGACCGCTCGGCTTATCCGTTAACTATCGTCTTG 5748
DB 4736 GTGTGACAGAACCCCGCTTACGCCGACCGCTCGGCTTATCCGTTAACTATCGTCTTG 4795
QY 5749 AGTCCAAACCCGGTAAAGACGACTTATCGCCACTGGGAGAGCCACTGGTAAACAGGATTA 5808
DB 4796 AGTCCAAACCCGGTAAAGACGACTTATCGCCACTGGGAGAGCCACTGGTAAACAGGATTA 4855
QY 5809 GCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGGCTTAACCTACGCT 5868
DB 4856 GCAGAGCAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTGGGCTTAACCTACGCT 4915
QY 5869 AACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAATA 5928

4916 A C A C T A G A A G G C A G C A G T A T T T G G T A T C T G C G C T C T G T G A A G C C A G T T A C C T T C G G A A A A 4975
 QY
 5929 G A G T T G G T A G C T C T T C A T C C G C A A A C A C C A C C G C T G T A G C G T G T T T T T T T T T 5988
 Db
 4976 G A G T T G G T A G C T C T T G A T C C G G C A A A C A A C C C G C T G T A G C G T G T T T T T T T T T 5035
 QY
 5989 G C A A G C A G C A G A T T A C G C G C A G A A A A A A A A G G A T C T C A G A A A G A T C C T T T G A T C T T T C T A 6048
 Db
 5036 G C A A G C A G C A G A T T A C G C G C A G A A A A A A A A G G A T C T C A G A A G A T C C T T T G A T C T T T C T A 5095
 QY
 6049 C G G G G T C T G A C G C T A G T G G A A C G A A A A C T A C A G T T A A G G A T T T T G T C A T G A G A T T A T 6108
 Db
 5096 C G G G G T C T G A C G C T A G T G G A A C G A A A A C T C A C G T T A A G G A T T T T G T C A T G A G A T T A T 5155
 QY
 6109 C A A A A G G A T C T C A C C T A G A T C C T T T T A A A T T A A A A T G A A G T T T T T T T T T T T T T T T 6165
 Db
 5156 C A A A A G G A T C T T C A C C T A G A T C C T T T T A A A T T A A A A A T G A A G T T T G C G C A A T C A A T C T 5215
 QY
 6166 A A A G T A T A T A G T A A A A C T T G G T C T G A C A G T T A C C A A T G C T T A A T C A G T G A G G C A C T A 6225
 Db
 5216 A A A G T A T A T A G T A A A C T T G G T C T G A C A G T T A C C A A T G C T T A A T C A G T G A G G C A C T A 5275
 QY
 6226 T C T A C G C A T C T G T C T A T T T C G T T C A T C A T A G T T G C C T G A C T C C C C G T C G T G T A G A T A A 6285
 Db
 5276 T C T A C G C A T C T G T C T A T T T C G T T C A T C C A T A G T T G C C T G A C T C C C C G T C G T G T A G A T A A 5335
 QY
 6286 C T A C G A T A C G G A G G C T T A C C A T C T G G C C C A G T G C T G C A T G A T A C C G C A G A C C C A C 6345
 Db
 5336 C T A C G A T A C G G A G G C T T A C C A T C T G G C C C A G T G C T G C A A T G A T A C C G C A G A C C C A C 5395
 QY
 6346 G C T C A C C G C T C C A G A T T A T C A G C A A T A A C A C G C C A G G A G G C C G A G C C A G A A 6405
 Db
 5396 G C T C A C C G C T C C A G A T T A T C A G C A A T A A C A C G C C A G G A G G C C G A G C C A G A A 5455
 QY
 6406 G T G G T C C T C A A C T T T A T C C G C C T C C A T C C A G T C T A T T A A T T G T T G C C G G A A G C T A G A G 6465
 Db
 5456 G T G G T C C T C A A C T T T A T C C G C C T C C A T C C A G T C T A T T A A T T G T T G C C G G A A G C T A G A G 5515
 QY
 6466 T A A G T A G T T G C C A G T T A T A G T T T G C G A A G T T G T T G C C A T T G C T A C A G C G - T C G T G G 6524
 Db
 5516 T A A G T A G T T G C C A G T T A T A G T T T G C G A A G T T T G C C A T T G T T G C C A T T G C T A C A G C C A T C G T G G 5575
 QY
 6525 T G T C A C G C T C G T C G T T G T A G G C T T C A T T C A G C T C C G T T C C C A A A G C A A C A T C A A G C G A G 6584
 Db
 5576 T G T C A C G C T C G T C G T T G T A T A G C T T C A T T A G C T C C G G T T C C C A A C A T C A A G C G A G 5635
 QY
 6585 T T A C A T G A T C C C C A T G T T G T G C A A A A A A A A G C G T T A G C T C C T T C G G T C C T C C G A T C G T T G 6644
 Db
 5636 T T A C A T G A T C C C C A T G T T G T G C A A A A A A A A G C G T T A G C T C C T T C G G T C C T C C G A T C G T T G 5695
 QY
 6645 T C A A A G T A A G T T G C C G C A G T T T A T C A T C A T G T T A T G G C A G C A C T G C A T A A T T C T 6704
 Db
 5696 T C A A A G T A A G T T G C C G C A G T T T A T C A T C A T G T T A T G C A G C A C T G C A T A A T T C T 5755
 QY
 6705 T T A C T G T A T C C A T C C G T A A G A T G C T T T T C T G T A C T G T A G T A C T C A A C C A A G T C A T 6764
 Db
 5756 T T A C T G T A T C C A T C C G T A A G A T G C T T T T C T G T A C T G T A G T A C T C A A C A A G T C A T 5815
 QY
 6765 T C T G A A T A G T A T G C G C G A C C A G T T G C T T T T G C C G C G T C C A A T A C C G G A T A A T A 6824
 Db
 5816 T C T G A A T A G T A T G C G C G A C C A G T T G C T T T G C C G C G T C A A C A C G G A T A A T A 5875
 QY
 6825 C C G C G C A C A T A G C A A A C T T T A A A A G T C T A T C A T T T G A A A A A C G T T C T T C G G G C G A A 6884
 Db
 5876 C C G C G C A C A T A G C A A A C T T T A A A A G T C T A T C A T T T G A A A A C G T T C T T C G G G C G A A 5935
 QY
 6885 A A C T C T C A G A G T T T A C C G T T G A G A T C C A G T T C G A T T A A C C A C T C G T G C A C C A 6944
 Db
 5936 A A C T C T C A G A G T T T A C C G T T G A G A T C C A G T T C G A T T A A C C A C T C G T G C A C C A 5995
 QY
 6945 A C T G A T C T T C A G C A T C T T T A C T T T C A C C A G G G T T T C T G G T G A G A A A A C A G G A A G C 7004
 Db
 5996 A C T G A T C T T C A G C A T C T T T A C T T T C A C C A G G G T T T C T G G T G A G A A A A C A G G A A G C 6055

QY 7005 A A A A T G C C G C A A A A A A A A G G A T A A G G C G A C A C G A A A T G T T G A A T A C T A T C T A C T T C C 7064
 Db 6056 A A A A T G C C G C A A A A A A A A G G A T A A G G C G A C A C G A A A T G T T G A A T A C T A T C T A C T T C C 6115
 QY 7065 T T T T T C A A T A T T A T T A A G C A T T T A T C A G G G T T A T T G T C T C A T G A C G G A T A C A T A T T T G 7124
 Db 6116 T T T T T C A A T A T T A T T A A G C A T T T A T C A - G G T T A T T G T C T A T G A C G G A T A C A T A T T T G 6174
 QY 7125 A A T G T A T T A G A A A A T A A A C A A A T A G G G T T C C G C G C A C A T T T C 7169
 Db 6175 A A T G T A T T A G A A A A T A A C A A A T A G G G T T C C G C G C A C A T T T C 6219

RESULT 14

US-09-826-025-1
 ; Sequence 1, Application US/09826025
 ; Patent No. US20020162123A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Lung-Ji
 ; TITLE OF INVENTION: Combination Immunogene Therapy
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/826,025
 ; FILING DATE: 04-Apr-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/838,702
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ingolia, Diane E.
 ; REGISTRATION NUMBER: 40,027
 ; REFERENCE/DOCKET NUMBER: CHANG-02687
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6145 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-826-025-1

Query Match 40.0%; Score 3062.2; DB 9; Length 6145;
 Best Local Similarity 71.7%; Pred. No. 0;
 Matches 4818; Conservative 0; Mismatches 1168; Indels 735; Gaps 32;
 QY 549 T T T T G A A A G A C C C C A C C T G T A G G T T T G C A A G C T A G C T T A A G T A A C G C C A C T T T G C A A G G 608
 Db 144 T T T T G A A A G A C C C C C C C G T A G G - T G C A A G C T A G C T T A A G T A A C G C C A C T T T G C A A G G 201
 QY 609 C A T G G - A A A A T A C A A C T G A A T A G A G A G T T C A G A T C A A G G T T A G A A C A G A G A C 667
 Db 202 C A T G A A A A A T A C A T A A C T A G A A A A A G T T C A G A T C A A G G T C A G A A A A A C 261
 QY 668 A G C A G A T A T G G C C C A A A C A G A T A T C T G T A A G A G C A G T T C C T G C C C C - G T C A G G G C C 726
 Db 262 A G C T G A A T A - - - - C C A A C A G A T A T C T G T A A G C G G T T C C T G C C C C G G C T C A G G G C C 317

Qy 727 AAGAACAGTTGGAACAGAGAGAAAT-ATGGGCCAAAACAGGATATCTGTGTGAAGCAGTTCT 785
 Db |||||
 Qy 318 AAGAACAGATGAGACAGCTGATGAGTGGGCCAAAACAGGATATCTGTGTGAAGCAGTTCT 377
 Db |||||
 Qy 786 GCCCGCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCCGCCTCAGCAGTTTC 845
 Db |||||
 Qy 378 GCCCGCGCTCAGGGCCAAAGAACAGATGGTCCCAAGATGGGTCCCGCCTCAGCAGTTTC 437
 Db |||||
 Qy 846 TAGAAGAACCATCAGATGTTTCCAGGGTGCCTCAAGGACCTG-AAATGACCTGTGCCTTA 904
 Db |||||
 Qy 438 TAGTGAATCATCAGATGTTTCCAGGGTGCCTCAAGGACCTGAATGATGACCTGTACCTTA 497
 Db |||||
 Qy 905 TTTGAATTAACCAATCAGTTGCTTCTGCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 964
 Db |||||
 Qy 498 TTTGAATTAACCAATCAGTTGCTTCTGCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 557
 Db |||||
 Qy 965 CAATAAAGAGCCCAACACCTCACTCGGGCGCCAGTCTCCGATAGATCGGTGCC 1024
 Db |||||
 Qy 558 CAATAAAGAGCCCAACACCTCACTCGGGCGCCAGTCTCCGATAGATCGGTGCC 617
 Db |||||
 Qy 1025 CGGTACCCGCTATCCCAATAAGCCTCTTGTGTTGTCATCCGAATCGTGACTCGCTG 1084
 Db |||||
 Qy 618 CGGTACCCGCTATCCCAATAAGCCTCTTGTGTTGTCATCCGAATCGTGACTCGCTG 677
 Db |||||
 Qy 1085 ATCTTGGGAGGCTCTCTCAGATTTGATGACTGCCACCTCGGGGTCTTTTCAATTTGGA 1144
 Db |||||
 Qy 678 TTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCCACGAGCGGGTCTTTTCAATTTGGG 737
 Db |||||
 Qy 1145 GGTTCACCGAGATTTGAGACCCCTGCGAGGACCAACCGACCCCGCGCGGAGGTA 1204
 Db |||||
 Qy 738 GGCTCGTCCGGATTTGAGACCCCTGCGCAGGACCAACCGA-CCCACACCGCGGAGGTA 796
 Db |||||
 Qy 1205 AGCTGGCCAGCAACTTATCTGTGCTGTCGATTGCTAGTGTCTATGACTGATTTTATG 1264
 Db |||||
 Qy 797 AGCTGGCCAGCAACTTATCTGTGCTGTCGATTGCTAGTGTCTATGACTGATTTTATG 856
 Db |||||
 Qy 1265 CGCTGCTCGGTACTAGTTAGCTAACTAGTCTGTATCTGCGGACCCGCTGGTGGAACT 1324
 Db |||||
 Qy 857 CGCTGCTCGTACTAGTTAGCTAACTAGTCTGTATCTGCGGACCCGCTGGTGGAACT 916
 Db |||||
 Qy 1325 GACGAGTTCGNAACACCCGCGGACCCCTGGGAGACGTCCTCAGGACTTCGGGGCGGT 1384
 Db |||||
 Qy 917 GACGAGTTCGNAACACCCGCGGACCCCTGGGAGACGTCCTCAGGACTTCGGGGCGGT 976
 Db |||||
 Qy 1385 TTTTGTGCGCCGACTGAGTCTTAAATCCCGATCGTTTAGGACTCTTTTGTGCAACCCC 1444
 Db |||||
 Qy 977 TTTTGTGCGCCGACTGAGGAGGAGTCTGATGGAATCCGACCC----- 1022
 Db |||||
 Qy 1445 CTTAGAGGAGGATATGTTGTTCTGGTAGGAGACGAGAACCTAAACAGTTTCCCGCTCC 1504
 Db |||||
 Qy 1023 -----CGTCAGGATATGTTGTTCTGGTAGGAGACGAGAACCTAAACAGTTTCCCGCTCC 1077
 Db |||||
 Qy 1505 GTCTGAATTTTGTCTTGGTTTGGGACCGAGCGCGCGCTCTGTGTTCTGTTCTGTTCTGTTCT 1561
 Db |||||
 Qy 1078 GTCTGAATTTTGTCTTGGTTTGGGACCGAGCGCGCTCTGTGTTCTGTTCTGTTCTGTTCTGTTCT 1137
 Db |||||
 Qy 1562 CAGCATCGTCTGTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1621
 Db |||||
 Qy 1138 CAGCATCGTCTGTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1197
 Db |||||
 Qy 1622 CCGGGCTAGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACTGGAAGAGTTCGA 1681
 Db |||||
 Qy 1198 CC-----AGACTGTTTACCACTCCCTTAAGTTTGAACCTTAGGTCACTGGAAGAGTTCGA 1251
 Db |||||
 Qy 1682 CGGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGAGCTGGGGTTACTTCTGCTC 1741
 Db |||||
 Qy 1252 CGGATCGCTCAACACAGTCGGTAGATGTCAAGAGAGAGCTGGGGTTACTTCTGCTC 1311
 Db |||||
 Qy 1742 TGCAGATGGCCAACTTTTAACTGCGATGCGCGGAGACCGGACCTTTAACCGAGACT 1801
 Db |||||
 Qy 1312 TGCAGATGGCCAACTTTTAACTGCGATGCGCGGAGACCGGACCTTTAACCGAGACT 1371
 Db |||||

Qy 1802 CATCACCCAGGTTAAGATCAAGGTTCTTTTACCTGGCCGCGATGACACCCAGACAGGT 1861
 Db |||||
 Qy 1372 CATCACCCAGGTTAAGATCAAGGTTCTTTTACCTGGCCGCGATGACACCCAGACAGGT 1431
 Db |||||
 Qy 1862 CCCCTACATCTGACCTGGGAAGCTTGGCTTTTGAACCCGCTCCTGGGTCAAGCCCTT 1921
 Db |||||
 Qy 1432 CCCCTACATCTGACCTGGGAAGCTTGGCTTTTGAACCCGCTCCTGGGTCAAGCCCTT 1491
 Db |||||
 Qy 1922 TGTACACCTTAAGCCTCGCCCTCTCTTCTTCCATCCGCGCCGCTCTCTCCCCCTTTGAACC 1981
 Db |||||
 Qy 1492 TGTACACCTTAAGCCTCGCCCTCTCTTCTTCCATCCGCGCCGCTCTCTCCCCCTTTGAACC 1551
 Db |||||
 Qy 1982 TCCTGTTTGAACCCGCGCTCGATCTCTCTTTATTCAGCCCTCACTCTCTCTCTAGGCGC 2041
 Db |||||
 Qy 1552 TCCTGTTTGAACCCGCGCTCGATCTCTCTTTATTCAGCCCTCACTCTCTCTCTAGGCGC 1611
 Db |||||
 Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101
 Db |||||
 Qy 1612 C-----GGAAATTCGATCTGA 1627
 Db |||||
 Qy 2102 CCCTGACATGACAAAGAGTTACTAACAGCCCTCTCTCCAAAGTCACTTTACAGGCTTCTAC 2161
 Db |||||
 Qy 1628 TCAAGAGACAGGATGAGGATCGTTTTCGATGATTTGAACAAAGATGGATTGACGAGGTTT 1687
 Db |||||
 Qy 2162 TTAGTTCACGACGAGGCTTGGAGACCTTGGCGGCGCTTACCAAGAAACAACTGACCGA 2221
 Db |||||
 Qy 1688 TCCGGCCGCTTGG-----GTGAGAGGCTTATTCGGCTATGACTGGGACAA 1733
 Db |||||
 Qy 2222 CCGTGTGTACTCACTTACCCCTTACGAGTGGCGACACAGTGTGGTGGTCCGCGACACAGACT 2281
 Db |||||
 Qy 1734 CAGAAATTCGGCTCTCTGATTCGCGCGCTTTCGGCTGTGACGCGAGGGCGCCCGGTT 1793
 Db |||||
 Qy 2282 AAGAACTTAGAACTCGCTGGAAGAGGCTTTACACAGTCTCTGTGACACCCACCCGCGC 2341
 Db |||||
 Qy 1794 CTTTTTTCAGACCGGACTGTCCGGTCCCTGAATGAAGTGC----- 1836
 Db |||||
 Qy 2342 CTCAAAGTAGACGGCATGCGAGCTTGGATACACCCGCGCCACGTAAGAGGCTGCGGACCC 2401
 Db |||||
 Qy 1837 -----AGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGCGGCTCTCTTGGCGAGC 1891
 Db |||||
 Qy 2402 GGGGTTGGACCATCTCTAGACTGCCATGGGATGGAGCTGTATCATCTCTCTTCTTGGTAG 2461
 Db |||||
 Qy 1892 TGTCTCGAGCTGTCTCACTGAAGCGGAAGGACTGGCTGTCTATTGGGCGAAGTGC CGGG 1951
 Db |||||
 Qy 2462 CAACAGCTACAGGTGTCCACTCCGACATCCAGCTGACCCAGAGCCCAAGACGCTGAGCG 2521
 Db |||||
 Qy 1952 GCAGATCTCTGTCTCTCACTTGTCTTCCGCGAGAAAGTATCCATCATGGCTGATGC 2011
 Db |||||
 Qy 2522 CCAGCGTGGGTGACAGAGTGACCATCACCTGTAAAGCCAGTCAAGATGTGGGTACTTCTG 2581
 Db |||||
 Qy 2012 AATCGCGCGCTGCATACGCTTGTATCCGGCTACCTGCGCCATTTCGACCAACCAAGCAACA 2071
 Db |||||
 Qy 2582 TAGCTTGGTACCAGACAGCAGGTAAGGCTCAAAGCTGCTGATCTACTCTGAGCATCCA 2641
 Db |||||
 Qy 2072 TCGCATCGAGCGACGCTACTCGGATGGAAGCGGCTCTTGTGATCAGGATGATCTGGA 2131
 Db |||||
 Qy 2642 CCGGCGACACTGTGTGCCAAGCAGATTTCAGCGGTAGCGGTAGCGGTACCGACTTCACCT 2701
 Db |||||
 Qy 2132 CGAAGACATCAGGGGCTCGGCGCAGCGCAACTGT-----TCGCGAGC 2175
 Db |||||
 Qy 2702 TCACCATCAGCAGCCTCCAGCAGAGGACATCGCCACCTACTACTGCGAGCAATATAGCC 2761
 Db |||||
 Qy 2176 TCAAGGCGCGCATCGCCGACGCGGAGGATCTCGTCTGTGACCCATGGCGATGCCTGCTTGC 2235
 Db |||||
 Qy 2762 TCTATCGGTGTTCCGCCAAGGACCAAGGTGGAAATCAAACGAGGTGGTCTCAGGATCGG 2821
 Db |||||
 Qy 2236 CGAATATCATGGTGGAAATGGCGCTTTTCTGATTTCACTGACTGTGGCGGCTGGGTG 2295
 Db |||||
 Qy 2822 GTGGATCCGCTCTGTTGGCTCAGGATCGGAGTCCAACTGGTGGAGAGCGGTGAGGTG 2881
 Db |||||
 Qy 2296 TGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATTTGCTGAAGAGCTTGGCG 2355
 Db |||||
 Qy 2882 TTGTGCAACCTGGCGGCTCCTCGCGCTGTCTCTCGCATCTGGCTTCGATTTTCAACA 2941
 Db |||||

Db 2356 GCGAATGGGCTGACCGCTCTCTCGTCTTACGGTATCGCGCTCC----- 2402
 QY 2942 CATATTGGATGAGTGTGGGTGAGACAGCAGCAGCTGGAAAAAGTCTTGAGTGGATTGGAGAA 3001
 Db 2403 -----GATTCGACGCGCATCGCCTTCTATCGCCTTCTTGACGAG 2441
 QY 3002 TTCAATCAGATAGCAGTACGATTAACATATGTCGCGCTCTTAAAGGATAGATTACAAATAT 3061
 Db 2442 TTCTTCTGAGCGGACTCTGGGGTTCGAAATGACCGACCAAGGAGCCCAACCTGCCAT 2501
 QY 3062 CCGCAGAACAGCCCAAGAACACATTTGTTCTGCAATGGACAGCCTGAGACCCGGAAGACA 3121
 Db 2502 CACGAGATTTGATTTCCACCGCGCCTTCTATGAAGGTTGGCTTCGGAATCGTTTTCC 2561
 QY 3122 CCGGGTCTATTTTGTGCAAGCCTTACTTTCGGCTTCCCTGGTTTGTCTATTTGGGGCC 3181
 Db 2562 GGGACGCC-----GGCTGGATGATCCT 2583
 QY 3182 AAGGAGCCCGGTCACCGTCTCCAGTGTCAAGCCACACAGCGCGCGCGCGACAC 3241
 Db 2584 CAGCGCGGGATCTCATGCTGAGTTCTTGGCCACCCCGGCTCGATCCCTCGCGAG 2643
 QY 3242 CAACACCGCGGCCACCATCGGTGCGAGCCTGTCCTTGGCGCCAGAGGGGCTCGGC 3301
 Db 2644 TTGGTTCACTGCTGCGCTGAGGCTGACGACCTCG----- 2678
 QY 3302 CAGCGCGGGGCGGAGTGACACAGAGGGGCTGACATTCGCGCTGGATCCCAACTCT 3361
 Db 2679 -----CGAGTTCTACCGGCAAGTCAAAATCG 2705
 QY 3362 GTTACCTGCTGGATGAATCCTTCTCATATGTTGTCTATTCTCATGCTGTTGTCCTGA 3421
 Db 2706 TCGCATCCAGGA---AACACAGCGGCTATCCGCGCATCCATGCCCCCGAACTGCGAG 2762
 QY 3422 GAGTGAAGTTTACAGGAGCGAGAGCCCGCGGTACAGCAGAGGCGCAGAACCAAGCTCT 3481
 Db 2763 AGTGGGAGCAGATGCGCGCTTTGTCACCCGAGCGGACCTCTCGCGCTGATAC 2822
 QY 3482 ATAAACAGCTCAATCTAGACAGAGAGAGAGTACGATGTTTGGACAGAGACGTGGCC 3541
 Db 2823 AGAACGAATTTGCTGAGGATCTCATGAGTGTGCTT-----CCCGTTTTTC 2869
 QY 3542 GGGACCTCAGATGGGGGAAAGCCGAGAGAGAGAACCTCAGGAGGCTGTACAATG 3601
 Db 2870 CCGCTGAGTCACTCGTGGATGAGCGCTGCGCTGCGCTGCGGACGCGAG-----CTG 2925
 QY 3602 AACTGCAGAAAGATAAGATGGGGAGGCGCTACAGTGAGATTTGGGATGAAAGCGAGCGCC 3661
 Db 2926 CTCACCACTCGAGGCGGTGACGCTGACAGAGCGCGAGTGCAGAACTGC----- 2978
 QY 3662 GGAGGGGCAAGGGGACGATGGCTTTTACAGGCTCTAGTACAGCAGCACCAAGGACACT 3721
 Db 2979 -----TCCAAAGGGACCTCAAGGCTTTCCAGGGACACTAGGCTGACTCCATCGAGCCAG 3033
 QY 3722 ACGAGCCCTTACATGACGAGGCGCTGCCCCCTCGCTAACTCGAGCGGCGCGGATCCGG 3781
 Db 3034 TGTA-----GAGATAAGCTTATCG 3052
 QY 3782 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTTCAGTCAACA 3841
 Db 3053 ATTAGTCCAAATTTGTTAAAGACAGGATATCAATGGTCCAGGCTCTAGTTTTCAGTCAACA 3112
 QY 3842 ATATCAACAGCTGAAGCCTATAGATGACGAGCCATAGATAAAATAAAGATTTTATTTAG 3901
 Db 3113 ATATCAACAGCTGAAGCCTATAGATGACGAGCCATAGATAAAATAAAGATTTTATTTAG 3172
 QY 3902 TCTCCAGAAAGGGGGGAAAGAGCCCACTGTTAGTTTGGAGCTAGCTAGCTTAAGT 3961
 Db 3173 TCTCCAGAAAGGGGGGAAAGAGCCCACTGTTAGTTTGGAGCTAGCTAGCTTAAGT 3232
 QY 3962 AACGCCATTTGCAAGGCAATG-AAAAATACATAAAGTGAATAGAGAAAGTTCAGATCAAG 4020

Db 3233 AACGCCATTTTGAAGGCGATGGAAAAATACATAAAGTGAATAGAGAAAGTTCAGATCAAG 3292
 QY 4021 GTTAGGAACAG- GAGACAGCAGATATGGGCCAAACAGGATATCTGTGTTAAGCAGTTTC 4079
 Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGTTAAGCAGTTTC 3352
 QY 4080 GTGCCCC--GCTCAGGCGCAAGAACAGTTTGGAAACAGGAAATATGGGCCAAACAGGATATC 4138
 Db 3353 CTGCCCCGGCTCAGGGCCAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATC 3412
 QY 4139 TGTGTAAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCAAGATCGGPTC 4198
 Db 3413 TGTGTAAAGCAGTTCTCTGCCCGCTCAGGGCCAAAGAACAGATGTTCCCAAGATCGGPTC 3472
 QY 4199 CCGCCCTCAGCAGTTTCTAGAGAACATCAGATGTTTCCAGGGTGCCTCAAGCACTGAA 4258
 Db 3473 CAGCCCTCAGCAGTTTCTAGAGAACATCAGATGTTTCCAGGGTGCCTCAAGCACTGAA 3532
 QY 4259 ATGACCCCTGTGCTTATTTTGAACCAATCAGTTTCTGCTTCTGCTTCTGTTCCGCGC 4318
 Db 3533 ATGACCCCTGTGCTTATTTGAACCAATCAGTTTCTGCTTCTGCTTCTGTTCCGCGC 3592
 QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGCGCCAGTCTCTCC 4378
 Db 3593 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGGGGCGCAGTCTCTCC 3652
 QY 4379 GATACATCGCTGCGCGGGTACCCGTTTCTCAATAAACCCCTCTTGCAGTTGCAATCGA 4438
 Db 3653 GATTGACTGAGTGCAGCGGATACCCGTTATCAATAAACCCCTCTTGCAGTTGCAATCGA 3712
 QY 4439 CTCGTGCTGCTGCTGCTTCTTGGGAGGCTCT--CTCTGAGTGAATGACTACCCGTCAGCGG 4497
 Db 3713 CTTGTGCTGCTGCTGCTTCTTGGGAGGCTCTCTCTGAGTGAATGACTACCCGTCAGCG 3772
 QY 4498 GGTCTTTCAGTTTCTCCACCTTACACAGCTCTCACTAAATTCCTGATGTGCGCGAGGA 4557
 Db 3773 GGGTCTTTCATTTGGGGGCTGCTCGGATCGGAGACCCCTGCGCCAGGACCAACCGACC 3832
 QY 4558 CTCCTGACGCGGTTTTTGTGTTTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617
 Db 3833 CACCAACCGGAGGTAAGCTG----- 3852
 QY 4618 ACTACATCTGACTCTCGGCTTTATAAAGAAATGTTGAAGGCTCTGTGGACTATCTGCC 4677
 Db 3853 -----GCTGCTCGCGGTTTCCGTTGATGACGTTGAACCTCTGAC-ACATGAGCT 3904
 QY 4678 ACAGACTTTTTTAAGATTTTATGCTCTCTGATGAGGATTTAGTCAATCTATCTCTCT 4737
 Db 3905 CCGGAGAGCGGTACAGCTTGTCTGTAAGCGATGCGGGAGCAGACAAAGCCGTCAGGG 3964
 QY 4738 CTAATTTGCTGCTCTCGTATTTTAAATTTTCTAGTTTGCATCCCTTCTGAGAGCAC 4797
 Db 3965 CCGCTCAGCGGTTGTTGGC----- 3984
 QY 4798 GSCGATGACAGATGTTAATACTCTGAGGCGAGCTTCTGTGAAAAGTTGCTGCTGGGCT 4857
 Db 3985 ----- 3984
 QY 4858 CAGTGTGAGATTTTGGCATAAAGGGGTCTGCGCCCTGTGTACAGACAGATCGGAATCT 4917
 Db 3985 ----- 3984
 QY 4918 AGAGTGCATATCAGAGTCCCGCGGTTCCGGGGCTCTGATCTCAGGGCATCTTTGCTTA 4977
 Db 3985 -----GGTCTCGGGGCGAGCCATGACCCAGTC----- 4012
 QY 4978 GAGATCCTCTAGCGCGGACGATCGTGGCGGTACCGAGCTCGAATTCGTAATCATGCT 5037
 Db 4013 -----ACGTAGCATAGCGAGTGTATATCTGGCTTAATCTATGCGCAT 4055
 QY 5038 CATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAAATTCACACAAACATACGAGCGG 5097
 Db 4056 CAGACAGATTTGACTGAGAG----- 4076

Db 1492 TGTACACCTTAAGCCTCGCCCTCTCTCTCTCCATCCGCCCGCTCTCTCCCCCTTGAACC 1551
Qy 1982 TCCTGTTTCGACCCCGCTCGATCTCTCTCTTATCCAGCCCTCACTCTCTCTAGGCGC 2041
Db 1552 TCCTGTTTCGACCCCGCTCGATCTCTCTCTTATCCAGCCCTCACTCTCTCTAGGCGC 1611
Qy 2042 CCCCATATGGCCATATGAGATCTTATATGGGGCACCCCGCCCTTGTAAACTTCCCTGA 2101
Db 1612 C-----GGAAATTCCGATCTGA 1627
Qy 2102 CCCTGACATGACAAGAGTTACTTAACAGACCCCTCTCTCAAGCTCACTTACAGGCTTCTAC 2161
Db 1628 TCAAGAGACAGGATGAGATGCTTTCGATGATTTGAACAAGATGATTCACGAGGTTTC 1687
Qy 2162 TTAGTCCACGACGAGTCTGGAGACCTCTGGCGGAGCTTACCAAGAAACAATCGACCGA 2221
Db 1688 TCCGCGCGCTGG-----GTGGAGAGGCTATTTCGGCTATGACTGGGCACAA 1733
Qy 2222 CGGTGGTACCTCACCTTACCGAGTCCGGGACACAGTGTGGGTCCGCGGACACAGACT 2281
Db 1734 CAGACAATCGGCTGCTCTGATCCCGCTGTTCGGCTGTTCAGCGCAGGGCGGCCGGTT 1793
Qy 2282 AAGAACCTAGAACCTCGTGGGAAGGACCTTACACAGTCTCTGCTGACACCCCCACCGCC 2341
Db 1794 CTTTGTGTCAGACCGACCTGTTCGGCTGCTGTAATGAATGCT----- 1836
Qy 2342 CTCAAAGTAGACGGCATCCGAGCTTGGATACACGCCGCCACGCTGAAGGCTGCCGACCCC 2401
Db 1837 -----AGACGAGGACGCGGCTATCGTGTGCTGCGCACGAGCGGCTTCTTGGCAGC 1891
Qy 2402 GGGGTGACCATCTCTTAGACTGCTCATGGAGTGGAGCTGTATCATCTCTTCTTGGTAG 2461
Db 1892 TGTGCTGACCTGTCTACTGAAGCGGAAGGAGCTGGCTGTATTGGCGGAAGTGC CGGG 1951
Qy 2462 CAAAGCTACAGTGTCTCACTCCGACATCCAGTCTGACCCAGGACCCAGAGCTGACGG 2521
Db 1952 GCAGATCTCTGTCTCATCTCACTTGTCTTCTGCTGCGAGAAATATCCATCATGCTGATGC 2011
Qy 2522 CCAGCTGGGTGACAGAGTGACATCACTCTGTAGCGCCAGTCAAGGATGGGTACTTCTG 2581
Db 2012 AATGGCGGCTGCATAGCTGTATCCGGCTACCTGCGCCATTCGACCAAGCGAACA 2071
Qy 2582 TAGCTTGGTACAGCAGAGCCAGGTAAAGGTCCAAAGCTCTGATCTACTTGGACATCCA 2641
Db 2072 TCGCATCGAGCGAGCAGCTACTCGATGGAAGCGGCTCTTGTGATCAGGATGATCTGA 2131
Qy 2642 CCGGCACACTGTGTGTCACAGCAGATTCAGGGTAGCGGTAGCGGTACCACTTCACT 2701
Db 2132 CGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGT-----TCGCCAGGC 2175
Qy 2702 TCACCATCAGCAGCTCCAGCCAGAGGACATCGCCACCTACTACTGCGCAGCAATATAGCC 2761
Db 2176 TCAAGCGCGCATGCCCGACGGGAGGATCTGCTGTGACCCATGCGATGCGCTGCTTGC 2235
Qy 2762 TCTATCGTCTGTGGCCCAAGGACCAAGGTGGAATCAACAGAGGTGGCTCAGGATCGG 2821
Db 2236 CGAATATCATGTGGAATGSCCGCTTTCTGGATTTCATGACTGTGGCGGCTGGGTG 2295
Qy 2822 TGTGATCCGGCTCTGGTGGCTCAGATCGGAGGTCCAACTGGTGAGAGCGGTGAGGTG 2881
Db 2296 TGGCGGACCGCTATCAGGACATAGCGTTGGGTACCCGTTGATATTGCTGAAGAGCTTGGCG 2355
Qy 2882 TTGTGCACTTGGCGGCTCCGCGCTGCTGCTCGCATCTGGCTTCGATTTCAACA 2941
Db 2356 GCGAATGGGCTGACCGCTTCTCTGCTTTCAGGTATCGCGCTCC----- 2402
Qy 2942 CATATTGATGAGTTGGGTGAGACAGGACCTGGAAAGGCTTTGAGTGGATTGGAGAAA 3001
Db 2403 -----GATTCGAGCGCATCGCTTCTATCGCTTCTTGAAG 2441
Qy 3002 TTATTCAGATAGCAGTACGATTAACATATGCGCGCTCTCTAAAGATAGATTTACAATAT 3061
Db 2442 TTCTTCTGAGCGGACTCTGGGGTTGGAATGACCGACCAAGCAGCGACGCCAACCTGCCAT 2501

Qy 3062 CGCGAGACAACGCCAAGACAACATTTGTTCTTSCAAATGGACAGCCTGAGACCCCGAAGACA 3121
Db 2502 CACGAGATTTGATTCACACCGCCCTCTATGAAGGTTGGGCTTCGAATCGTTTTCC 2561
Qy 3122 CCGGGGTCTATTTTGTGCAAGCCTTTACTTCGGCTTCCCTGGTTGTTTATTTGGGGCC 3181
Db 2562 GGGAGGCC-----GGCTGGATGATCCT 2583
Qy 3182 AAGGGACCCCGGTCAACCGTCTCCAGTCTTCCAGTCTTCCAGCCACCAAGCGCGCGGACAC 3241
Db 2584 CCAGCGGGGATCTCATGCTGGAGTTCTTCGCCCAACCCCGGGCTCGATCCCTCGCGAG 2643
Qy 3242 CAACACCGCGCCCAACCATCGCTGCGAGCCCTGTCTCCCTGCGCCAGAGCGGCTCGGC 3301
Db 2644 TTGGTTCACTGCTGCTCGCTGAGGCTGGACGACCTCG----- 2678
Qy 3302 CAGCGCGGGGGCGCAGTGCACAGAGGGGGCTGGACTTCCCTCTGATCCCAACTCT 3361
Db 2679 -----CGAGTTCTACCGGCGAGTGCAAATCCG 2705
Qy 3362 GCTACCTGCTGATGGAATCCTCTTCTATGCTGTCATCTCCTGCTTGTCTCTGA 3421
Db 2706 TCGCATCCAGGA---AACAGAGCGGCTATCCGGCATCCATGCCCCCGAAGTGCAGG 2762
Qy 3422 GAGTGAAGTTTCAGCAGGAGCGCAGAGCCCGCGGTACACAGCAGGGCCAGAACAGCTCT 3481
Db 2763 AGTGGGAGGACGATGGCGCTTTTGTGCGACCGGACGGGACGCTCTCGCGCTGATAC 2822
Qy 3482 ATAAAGAGCTCAATCTAGGAGCAAGAGAGGATGAGTGTGTTGGACAGAGAGCTGGCC 3541
Db 2823 AGAACGAATTTGCTTGCAGGCTCTCATGAGTGTCTT-----CCCGTTTTTC 2869
Qy 3542 GGGACCTGAGATGGGGGGAAGCCGAGAGGAAGAACCTCAGAGAGGCGCTGTACAATG 3601
Db 2870 CGCTGAGGTCATGCTGCTGATGAGCGCTGGCGCTGCTCGCGACGGCGAG---CTG 2925
Qy 3602 AACTGCAAGAAGATAAGATGGCGGAGCCTACAGTGAGATTGGGATGAAGCGCGAGCGCC 3661
Db 2926 CTCACCAACCCACTCGAGGGGCTGCGAGCGTGCAGAGCCGAGTGCAGAACTGC----- 2978
Qy 3662 GGAGGGGAGAGGCGCAGTGGCCCTTTACAGGCTCTCAGTACAGCCCAAGAGACCT 3721
Db 2979 -----TCCAAAGGACCTCAAGGCTTTCCGAGGACACTAGGCTGACTCCATCGAGCCAG 3033
Qy 3722 ACAGAGCCCTTCATGACGCGCCCTGCCCCCTGCTTAACCTGACGCGCGCGGATCCGG 3781
Db 3034 TGTA-----GAGATAAGCTTATCG 3052
Qy 3782 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACA 3841
Db 3053 ATTAGTCCAAATTTGTTAAAGACAGGATATCAGTGGTCCAGGCTCTAGTTTGTACTCAACA 3112
Qy 3842 ATATCAAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTAG 3901
Db 3113 ATATCAAGCTGAAGCTTATAGAGTACGAGCCATAGATAAATAAAGATTTTATTTAG 3172
Qy 3902 TCTCCAGAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGT 3961
Db 3173 TCTCCAGAAAGGGGGAATGAAGACCCACCTGTAGTTTGGCAAGCTAGCTTAAGT 3232
Qy 3962 AACCCCAATTTTGAAGGATCGG---AAATACATAACTAGATAAGAGTTCAGATCAAG 4020
Db 3233 AACCCCAATTTTGAAGGATCGGAAATAATACATACTAGATAAGAGTTCAGATCAAG 3292
Qy 4021 GTTAGGAACAGA--GAGACAGCAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTC 4079
Db 3293 GTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTC 3352
Qy 4080 CTGCCCC--GCTCAGGGCCAGAAACAGTTCGACAGGAGATATGGGCCAAACAGGATATC 4138
Db 3353 CTGCCCCGGCTCAGGGCCAGAAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATC 3412

QY 4139 TGTGTAAGCAGTTCTGTCGCCCGGTCTAGGGCCAGAAACAGATGTTCCCAAGATCGGTG 4198
 Db 3413 TGTGTAAGCAGTTCTGTCGCCCGGTCTAGGGCCAGAAACAGATGTTCCCAAGATCGGTG 3472
 QY 4199 CGCCCTCAGCAGTTCTAGAGAACATCAGATGTTCCAGGGTGCCTCAAGACCTGAA 4258
 Db 3473 CAGCCCTCAGCAGTTCTAGAGAACATCAGATGTTCCAGGGTGCCTCAAGACCTGAA 3532
 QY 4259 ATGACCTGTGCTTATTTGAATACCAATCAGTTGCTTCTCGTCTCTGTTTCGGCGC 4318
 Db 3533 ATGACCTGTGCTTATTTGAATACCAATCAGTTGCTTCTCGTCTCTGTTTCGGCGC 3592
 QY 4319 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGGCGCCAGTCTCC 4378
 Db 3593 TTCTGCTCCCGAGCTCAATAAAGAGCCCAACCCCTCACTCGCGGCGCCAGTCTCC 3652
 QY 4379 GATAGACTCGTGCCTGCGCGGTACCGGTCTCCTGATGATGACTACCCGTCAGCGG 4438
 Db 3653 GATAGACTCGTGCCTGCGCGGTACCGGTCTCCTGATGATGACTACCCGTCAGCGG 3712
 QY 4439 CTGCTGCTCGTGCCTGCGCGGTACCGGTCTCCTGATGATGACTACCCGTCAGCGG 4497
 Db 3713 CTGCTGCTCGTGCCTGCGCGGTACCGGTCTCCTGATGATGACTACCCGTCAGCGG 3772
 QY 4498 GGTCTTTTCAAGTTTCTCCCACTACACAGGTCTCACTAAATTCCTGATGTCGCGCAGGGA 4557
 Db 3773 GGTCTTTTCAAGTTTCTCCCACTACACAGGTCTCACTAAATTCCTGATGTCGCGCAGGGA 3832
 QY 4558 CTCGCTCAGCCCGGTTTTTGTATTAATAAATGCAAGAACAGTGTTCCTTCAAGCCAG 4617
 Db 3833 CACCACCGGAGTAACTG- 3852
 QY 4618 ACTACATCTGACTCTCGGCTTATAAAGAAATGTTGAAGGCTCTGTGGAATATCTGCC 4677
 Db 3853 -----GCTGCTCGCGGTTCGGTGATGACGGTGAAGAACTCTGAC-ACATGAGCT 3904
 QY 4678 ACACGACTTTTAAAGATTTTATGCTCCTCGATGAGGATTTAGTCAATCTATCTCGT 4737
 Db 3905 CCGGAGAGCTCAGCTTGTCTGAAGCGATGCGGGAGCAGACAGCCGTCAGGG 3964
 QY 4738 CTATTTGTGCTTCTCGTATTTTAAATTTCTAGTTGCACTCCCTTCTGAGAGCAC 4797
 Db 3965 CCGCTCAGCGGTGTGGCG- 3984
 QY 4798 GCGATTGACAGTAGTAACTCTGAGGCGAGGCTTCTGTGAAGGTTGCTTGGGT 4857
 Db 3985 ----- 3984
 QY 4858 CAGTGTGAGATTTTGCATATAAAGGGTCTGTCCTGTTGACAGACAGATCGGAATCT 4917
 Db 3985 ----- 3984
 QY 4918 AGAGTGCATCTCAGAGTCCCGGCTTCGGGCTCTGATCTCAGGGCATCTTTGCTTA 4977
 Db 3985 -----GCTGTCGGGGCGCAGCCATGACCCAGTC- 4012
 QY 4978 GAGATCCTCTAGCCGAGCAGTCTGTCGGCGGGTACCGAGCTCGAATCGTAATCGT 5037
 Db 4013 -----ACGTAGCGATAGCGAGTGTATCTAGGCTTAATCTATGCGGCAT 4055
 QY 5038 CATAGCTGTTCTGTCGAATTTGTTATCCGCTCACAATTCACACACATACGAGCGG 5097
 Db 4056 CAGACGAGTTGACTGAGAG- 4076
 QY 5098 GAAGCATAAAGTGTAAAGCTCGGGTGCCTAATGAGTGAGCTAACTCAATTAATTCGT 5157
 Db 4077 ----- 4076
 QY 5158 TGGCTCAGTGCCTGCTTCCAGTCGGGAAACCTGTGTCGACGCTGCAATTAATGATCG 5217
 Db 4077 -----TGACCATATGCTGTGAATAC 4100
 QY 5218 GCCAACGCGGGGAGAGCGGTTTCGTTATTTGGGGCTTCTCGGCTTCTCGCTCAGTCG 5277

Db 4101 CGCACAGATGCGTAAGGAGAAATAACCGCATCAGGCGCTCTTCCGCTTCTCCTCGCTCAGTCG 4160
 QY 5278 ACTGCTCGCTCGGTTCGGTTCGGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA 5337
 Db 4161 ACTGCTCGCTCGGTTCGGTTCGGCTGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAA 4220
 QY 5338 TACGCTTATCCACAGATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCCAGC 5397
 Db 4221 TACGCTTATCCACAGATCAGGGGATAACGAGGAAAGAAACATGTGAGCAAAAGGCCAGC 4280
 QY 5398 AAAAGCCAGGAAACCGTAAAAAGCGGTTCGCTGGCGTTTTTTCATAGGCTTCGCGCCCG 5457
 Db 4281 AAAAGCCAGGAAACCGTAAAAAGCGGTTCGCTGGCGTTTTTTCATAGGCTTCGCGCCCG 4340
 QY 5458 CTGACGAGCATCACAAAATCGACGCTCAAGTTCAGAGTGGGAAACCGCAGCAGACTAT 5517
 Db 4341 CTGACGAGCATCACAAAATCGACGCTCAAGTTCAGAGTGGGAAACCGCAGCAGACTAT 4400
 QY 5518 AAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCG 5577
 Db 4401 AAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCTGCG 4460
 QY 5578 GGTCTTACCGGATACCTGTCGCGTTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCT 5637
 Db 4461 GGTCTTACCGGATACCTGTCGCGTTTTCTCCCTTCGGAAGCGTGGCGTTTTCTCATAGCT 4520
 QY 5638 CAGCTGTAGGTATCTCACTGTTGCGGTAGTGTGCTTCCCAAGCTGGCTGTGTGCGACG 5697
 Db 4521 CAGCTGTAGGTATCTCACTGTTGCGGTAGTGTGCTTCCCAAGCTGGCTGTGTGCGACG 4580
 QY 5698 AACCCCCGTTTCAGCGCGACCGCTCGCTTATCCGTTAACTATCTGTTGAGTCCAAACC 5757
 Db 4581 AACCCCCGTTTCAGCGCGACCGCTCGCTTATCCGTTAACTATCTGTTGAGTCCAAACC 4640
 QY 5758 CGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAAGAGATTAGCAGAGCA 4700
 Db 4641 CGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAAGAGATTAGCAGAGCA 4700
 QY 5818 GGTATGTAGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAA 5877
 Db 4701 GGTATGTAGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAA 4760
 QY 5878 GACAGATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGGAAGAAAGAGTTGTTA 5937
 Db 4761 GACAGATTTTGGTATCTGCGCTCTGCTGAAGCAGTTACCTTCGGAAGAAAGAGTTGTTA 4820
 QY 5938 GCTCTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTTGTTCGAAAGCAGC 5997
 Db 4821 GCTCTTGATTCGGGCAAAACAAACCCGCTGGTAGCGGTGTTTTTTTGTTCGAAAGCAGC 4880
 QY 5998 AGATTACGGGCAAAAGAAAGATCTCAAGAGATCTTGTGATCTTTTCTACGGGTCTG 6057
 Db 4881 AGATTACGGGCAAAAGAAAGATCTCAAGAGATCTTGTGATCTTTTCTACGGGTCTG 4940
 QY 6058 ACGCTCAGTGAAACCAAACTCACCTTAAGGATTTTGTGTCATGAGATTATCAAAAGGA 6117
 Db 4941 ACGCTCAGTGAAACCAAACTCACCTTAAGGATTTTGTGTCATGAGATTATCAAAAGGA 5000
 QY 6118 TCTTCACTAGTCTCTTTTAAATTAAGTTTAAATCAATCAATCAATCAATCAATCAATCAAT 6177
 Db 5001 TCTTCACTAGTCTCTTTTAAATTAAGTTTAAATCAATCAATCAATCAATCAATCAATCAAT 5060
 QY 6178 AGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTCAGGATCT 6237
 Db 5061 AGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACTATCTCAGGATCT 5120
 QY 6238 GTCATTTGCTTATCCATAGTTGCTGATCTCCCGCTGCTGAGTAACTACGATACGGG 6297
 Db 5121 GTCATTTGCTTATCCATAGTTGCTGATCTCCCGCTGCTGAGTAACTACGATACGGG 5180
 QY 6298 AGGGCTTACCATCTCGGCCCGAGTGTGCAATGATACCGCGAGACCCACGCTCACCGGCTC 6357

Db	5181	AGGGCTTACCATCTGGCCCGAGTGTGCAATGATACCGCGAGACCCACGCTCACGGGCTC	5240
Qy	6358	CAGATTATACGAAATAAACAGCCAGCCGGAAGCCGCGCAGAGAGTGGTCTCTGCAA	6417
Db	5241	CAGATTATACGAAATAAACAGCCAGCCGGAAGCCGCGCAGAGAGTGGTCTCTGCAA	5300
Qy	6418	CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAAGTAGTTCGC	6477
Db	5301	CTTTATCCGCTCCATCCAGTCTATTAAATTTGTCGGGGAAGCTAGAGTAAGTAGTTCGC	5360
Qy	6478	CAGTTAATAGTTTGGCAACGTTGTTGCAATTTGCTTACAGGC-TGCTGGTGTCAAGCTCGT	6536
Db	5361	CAGTTAATAGTTTGGCAACGTTGTTGCAATTTGCTTACAGGCATCGTGGTGTCAAGCTCGT	5420
Qy	6537	CGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCC	6596
Db	5421	CGTTTGGTATGGCTTCAATTCAGCTCCGGTTCCTCAACGATCAAGCGAGTTACATGATCCC	5480
Qy	6597	CCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGT	6656
Db	5481	CCATGTTGTGCAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGTTGTGCAAGTAAGT	5540
Qy	6657	TGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAAATCTCTTACTCTCATGC	6716
Db	5541	TGGCGCAGTGTATCACTCATGTTATGGCAGCACTGCATAAATCTCTTACTCTCATGC	5600
Qy	6717	CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTTGAGAATAGT	6776
Db	5601	CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCAATCTTGAGAATAGT	5660
Qy	6777	GTATGCGGCGACCGAGTTGCTCTTGCCTCCGGCGTCAATAACGGGATAATACCGGCGACATA	6836
Db	5661	GTATGCGGCGACCGAGTTGCTCTTGCCTCCGGCGTCAATAACGGGATAATACCGGCGACATA	5720
Qy	6837	GCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTGGGGCGGAAACTCTCAAGGA	6896
Db	5721	GCAGAACTTTAAAGTGTCTCATCTGAAACGTTCTTGGGGCGGAAACTCTCAAGGA	5780
Qy	6897	TCTTACCGCTGTTGAGATCCAGTTCCGATGTAAACCCACTCGTGCACCCCACTGTCTTCAG	6956
Db	5781	TCTTACCGCTGTTGAGATCCAGTTCCGATGTAAACCCACTCGTGCACCCCACTGTCTTCAG	5840
Qy	6957	CATCTTTTACCTTTCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA	7016
Db	5841	CATCTTTTACCTTTCACGAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAA	5900
Qy	7017	AAAGGGGAATAAGGGCGACGGAATGTTGAATACTACTCATCTCTTCTTTTCAATATT	7076
Db	5901	AAAGGGGAATAAGGGCGACGGAATGTTGAATACTACTCATCTCTTCTTTTCAATATT	5960
Qy	7077	ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTTATTAGA	7136
Db	5961	ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTTATTAGA	6020
Qy	7137	AAATAAACAATAAGGGTTCGCGGCACATTTCCCGGAAAGTCCACCTGACGTCTAAG	7196
Db	6021	AAATAAACAATAAGGGTTCGCGGCACATTTCCCGGAAAGTCCACCTGACGTCTAAG	6080
Qy	7197	AAACCATTTATCATGACATTACCTATATAAATAGGCGTATCACGAGGCCCTTTCGTC	7256
Db	6081	AAACCATTTATCATGACATTACCTATATAAATAGGCGTATCACGAGGCCCTTTCGTC	6140
Qy	7257	T 7257	
Db	6141	T 6141	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 19:41:59 ; Search time 289 Seconds
(without alignments)
592.854 Million cell updates/sec

Title: US-10-006-771B-2

Perfect score: 2350

Sequence: 1 MGWSCIILFLVATATGVHSD.....LSTATKDTYDALHMQALPPR 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2350	100.0	443	6	ABG76488 Humanised
2	2350	100.0	443	6	ABG74240 Chimaeric
3	1590.5	67.7	473	2	AAW26646 Chimeric
4	1590.5	67.7	514	2	AAW26647 Chimeric
5	1447	61.6	532	2	AAW73051
6	1409.5	60.0	643	2	AAW73050
7	1406.5	59.9	651	2	AAW26649
8	1406.5	59.9	692	2	AAW26650
9	1304.5	55.5	643	2	AAW73048
10	1228	52.3	634	6	ABB82300
11	1198	51.0	633	3	AAW84965
12	1094	46.6	444	2	AAW36845
13	1038	44.2	443	2	AAW24025
14	1015.5	43.2	461	2	AAW85508
15	991	42.2	523	3	AAW44994
16	986	42.0	524	3	AAW44995
17	966	41.1	428	2	AAW24027
18	965	41.1	352	2	AAW06272
19	937.5	39.9	574	7	AAW52216
20	937.5	39.9	2606	7	AAW52217
21	937.5	39.9	2606	7	AAW52214
22	920	39.1	411	8	ADR46829
23	918	39.1	483	2	AAW82315
24	910	38.7	631	7	ADH34696
25	908	38.6	403	2	AAW26648

26	881.5	37.5	456	3	AAW44992	Ray44992 M798cFv-i
27	879.5	37.4	264	2	AAW73049	Aaw73049 Humanised
28	876	37.3	450	3	AAW44991	Ray44991 M798cFv-i
29	871.5	37.1	333	3	AAW70111	Ray70111 Anti-Her2
30	871	37.1	283	5	AAU75160	Aau75160 Kabat con
31	871	37.1	283	6	ABG73148	Abg73148 Kabat con
32	871	37.1	283	6	ABG73868	Abg73868 Kabat con
33	870	37.0	263	2	AAW97889	Aaw97889 Kabat con
34	865.5	36.8	247	8	ADO40408	Ado40408 Single ch
35	865.5	36.8	252	8	ADO40405	Ado40405 Single ch
36	865.5	36.8	253	8	ADO40406	Ado40406 Single ch
37	865.5	36.8	254	8	ADO40412	Ado40412 Single ch
38	862.5	36.7	252	8	ADO40409	Ado40409 Single ch
39	846.5	36.0	252	8	ADO58062	Ado58062 S2 cell d
40	846	36.0	239	8	ADO40413	Ado40413 Single ch
41	842.5	35.9	252	3	AAW28158	Aaw28158 Human ant
42	839.5	35.7	479	6	ABU09530	Abu09530 Hamster a
43	839.5	35.7	551	6	ABU09531	Abu09531 Hamster a
44	826.5	35.2	951	3	AAW44993	Ray44993 DC86cFv-e
45	824	35.1	304	2	AAW40070	Aaw40070 Human H11

ALIGNMENTS

RESULT 1
ABG76488
ID ABG76488 standard; protein; 443 AA.
XX
AC ABG76488;
XX
DT 13-MAY-2003 (first entry)
XX
DE Humanised anti-CEA antibody, hMN14.
XX
KW CEA; carcinoembryonic antigen; IgTCR; T-cell receptor; cancer; tumour;
KW colorectal cancer; breast cancer; lung cancer; hMN14; cytostatic; mouse;
KW human; zeta signalling chain; CD8alpha hinge; humanised antibody.
XX
OS Homo sapiens.
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX
PN US2002165360-A1.
XX
PD 07-NOV-2002.
XX
PF 10-DEC-2001; 2001US-00006771.
XX
PR 30-NOV-2000; 2000US-0250087P.
PR 30-NOV-2000; 2000US-0250090P.
XX
PA (JUNG/) JUNGHANS R P.
XX
PI Junghans RP;
XX
DR WPI; 2003-298705/29.
DR N-PSDB; ABX13168.
XX
CC New chimeric molecule from humanized antibody against carcinoembryonic antigen and having signaling molecules of T cells and other effector cells, useful for the treatment of colorectal, breast and lung cancers.
XX
PS Disclosure; Page 7-8; 20pp; English.
XX
CC The invention relates to a chimaeric molecule comprising the carcinoembryonic antigen (CEA) binding domain of humanised antibody MN14 as a single chain antibody with a (GGSGS)3 linker, the zeta signaling chain of the T cell receptor (TCR) and an intervening CD8alpha hinge in which the cysteine residues have been mutated, with the IgTCR molecule occupying nucleotides 2426-3766 of the retroviral vector sequence.
CC appearing as ABX13168. The new chimaeric molecule expressed in T cells,

CC NK (not defined) or other effector cells are useful in treating patients
 CC with cancers expressing the CEA antigen, together with other or with
 CC heterologous constructs to engage additional stimulatory and functional
 CC properties of the effector cells to enhance the anti-tumour therapeutic
 CC efficacy. The cancer disorder includes colorectal, breast and lung
 CC cancers. The present sequence represents the chimaeric molecule of the
 CC invention
 XX
 XX SQ Sequence 443 AA;
 Query Match 100.0%; Score 2350; DB 6; Length 443;
 Best Local Similarity 100.0%; Pred. No. 1.3e-154;
 Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60
 DB 1 MGWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60
 QY 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDFTTISSLOPEDIATYCCQYSLYRSFGQGT 120
 DB 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDFTTISSLOPEDIATYCCQYSLYRSFGQGT 120
 QY 121 KVEIKRGGSGSGSGSGSEVQLVESGGGVQVQPSRLRLSCSASGDFDTTYMMSWRQ 180
 DB 121 KVEIKRGGSGSGSGSGSEVQLVESGGGVQVQPSRLRLSCSASGDFDTTYMMSWRQ 180
 QY 181 APGKLEWIGETHPDSSITINAPSLKDRFTISRDNKNTLFLQDLSLRPEDTGVYFCASL 240
 DB 181 APGKLEWIGETHPDSSITINAPSLKDRFTISRDNKNTLFLQDLSLRPEDTGVYFCASL 240
 QY 241 YFGFPWFAYWGQGTPTVTSVSSAKPTTTPAPRPTPTPTIASQPLSLRPAARPAAGAVHT 300
 DB 241 YFGFPWFAYWGQGTPTVTSVSSAKPTTTPAPRPTPTPTIASQPLSLRPAARPAAGAVHT 300
 QY 301 RGLDFALDPKLCYLLDGLFIYGVILTALFLRVKFSRAEPPAYQOQGNQLYNELNLGR 360
 DB 301 RGLDFALDPKLCYLLDGLFIYGVILTALFLRVKFSRAEPPAYQOQGNQLYNELNLGR 360
 QY 361 EYDVLDKRRGRDPEMGKPRKPNQEGLYNELQKMAEYSEIGMKGERRRGKHDGL 420
 DB 361 EYDVLDKRRGRDPEMGKPRKPNQEGLYNELQKMAEYSEIGMKGERRRGKHDGL 420
 QY 421 YQGLSTATKDTYDALHMQALPPR 443
 DB 421 YQGLSTATKDTYDALHMQALPPR 443
 RESULT 2
 ABG74240
 ID ABG74240 standard; protein; 443 AA.
 AC ABG74240;
 XX
 XX 23-OCT-2003 (revised)
 DT 22-APR-2003 (first entry)
 XX
 DE Chimaeric hMN14/T-cell receptor.
 XX
 XX Retroviral vector; T-cell receptor; hMN14; antibody; IgTCR; receptor;
 KW cytostatic; dermatological; neuroprotective; immunostimulant; GD3;
 KW ganglioside antigen; MB3.6; PSMA; tumour; 3D8; 4D4; 3E11;
 KW prostate-specific membrane antigen; zeta signalling chain; human; cancer;
 KW melanoma; neuroendocrine tumour; prostate cancer; small cell lung cancer;
 KW mouse; CD8alpha hinge.
 XX
 XX Homo sapiens.
 OS Mus sp.
 OS Chimeric.
 XX
 XX US2002132983-A1.
 FN
 XX
 XX 19-SEP-2002.
 PD
 XX

PF 10-DEC-2001; 2001US-00006773.
 XX
 XX 30-NOV-2000; 2000US-0250087P.
 PR 30-NOV-2000; 2000US-0250089P.
 XX
 XX (JUNG/) JUNGHANS R P.
 PA
 XX Junghans RP;
 PI
 XX
 XX WPI; 2003-208946/20.
 DR N-PSDB; ABX16565.
 XX
 PT New chimeric molecule useful in treating patients with disorders, such as
 PT melanoma, neuroendocrine disorders, prostate and small cell lung cancer
 PT comprises GD3 and/or PSMA binding domains of antibody.
 XX
 XX Disclosure; Page 7-8; 35pp; English.
 XX
 CC The invention relates to a chimaeric molecule comprising the GD3
 CC (ganglioside antigen) binding domain of antibody MB3.6, with any of 3
 CC variable gene sequences, or the PSMA (prostate-specific membrane antigen)
 CC binding domain of antibody 3D8, 4D4 and 3E11, with variable gene
 CC sequences, the zeta signalling chain of the T cell receptor and an
 CC intervening CD8alpha hinge in which cysteine residues have been mutated.
 CC The chimaeric molecules expressed in T cells or NK cells or other
 CC effector cells are useful in treating patients with cancers expressing
 CC the GD3 (MB3.6 derivatives) or PSMA antigen (3D8, 4D4, 3E11 derivatives),
 CC and/or together with each other or with heterologous constructs to engage
 CC additional stimulatory and functional properties of the effector cells to
 CC enhance the antitumour therapeutic efficacy (claimed). They are
 CC particularly useful in disorders including melanoma, neuroendocrine
 CC tumours and prostate and small cell lung cancer. The present sequence is
 CC a hMN14 antibody (specific to CEA antigen) in a fusion protein with the
 CC modified CD8alpha hinge and the T-cell receptor zeta chain (IgTCR)
 CC encoded by a retroviral vector. The hMN14 antibody coding region is
 CC replaced with the MB3.6, 3D8, 4D4 or 3E11 genes of the invention.
 CC (Updated on 23-OCT-2003 to standardise OS field)
 XX
 XX Sequence 443 AA;
 SQ
 Query Match 100.0%; Score 2350; DB 6; Length 443;
 Best Local Similarity 100.0%; Pred. No. 1.3e-154;
 Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60
 DB 1 MGWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTITCKASQDVGTSVAMVYQKPG 60
 QY 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDFTTISSLOPEDIATYCCQYSLYRSFGQGT 120
 DB 61 KAPKLLIYWTSTRTHTGVPSRFGSGSGTDFTTISSLOPEDIATYCCQYSLYRSFGQGT 120
 QY 121 KVEIKRGGSGSGSGSGSEVQLVESGGGVQVQPSRLRLSCSASGDFDTTYMMSWRQ 180
 DB 121 KVEIKRGGSGSGSGSGSEVQLVESGGGVQVQPSRLRLSCSASGDFDTTYMMSWRQ 180
 QY 181 APGKLEWIGETHPDSSITINAPSLKDRFTISRDNKNTLFLQDLSLRPEDTGVYFCASL 240
 DB 181 APGKLEWIGETHPDSSITINAPSLKDRFTISRDNKNTLFLQDLSLRPEDTGVYFCASL 240
 QY 241 YFGFPWFAYWGQGTPTVTSVSSAKPTTTPAPRPTPTPTIASQPLSLRPAARPAAGAVHT 300
 DB 241 YFGFPWFAYWGQGTPTVTSVSSAKPTTTPAPRPTPTPTIASQPLSLRPAARPAAGAVHT 300
 QY 301 RGLDFALDPKLCYLLDGLFIYGVILTALFLRVKFSRAEPPAYQOQGNQLYNELNLGR 360
 DB 301 RGLDFALDPKLCYLLDGLFIYGVILTALFLRVKFSRAEPPAYQOQGNQLYNELNLGR 360
 QY 361 EYDVLDKRRGRDPEMGKPRKPNQEGLYNELQKMAEYSEIGMKGERRRGKHDGL 420
 DB 361 EYDVLDKRRGRDPEMGKPRKPNQEGLYNELQKMAEYSEIGMKGERRRGKHDGL 420
 QY 421 YQGLSTATKDTYDALHMQALPPR 443

DB 421 YQGLSTATKOTYDALHMQALPPR 443

|||||

RESULT 3

AAW26646

ID AAW26646 standard; protein; 473 AA.

XX AAW26646;

XX

DT 17-OCT-2003 (revised)

DT 12-FEB-1998 (first entry)

XX

XX Chimeric receptor hCTMO1/CD8/zeta.

DE

XX Cell activation; chimeric receptor; DNA delivery; CTMO1; scFv; CD8;

KW T cell receptor zeta chain; cell proliferation; cytokine; inflammation;

KW effector; cell differentiation; antibody secretion; phagocytosis;

KW tumour infiltration; adhesion; infection; cancer; allergy;

KW rheumatoid arthritis; osteoarthritis; asthma; eczema;

KW inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;

KW psoriasis; multiple sclerosis; transplant rejection; diabetes;

KW graft versus host disease; human; therapy.

XX

OS Homo sapiens.

OS Chimeric.

XX

PN WO9723613-A2.

XX

PD 03-JUL-1997.

XX

XX 23-DEC-1996; 96WO-GB003209.

XX

XX 21-DEC-1995; 95GB-00026131.

XX

PA (CLLT) CELLTech THERAPEUTICS LTD.

XX

XX Bebbington CR, Lawson ADG, Weir ANC, Finney HM;

XX

XX WPI; 1997-351052/32.

DR N-PSDB; AAT90509.

DR

XX

PT New DNA systems for activating cells - comprising DNA coding for a

PT chimeric receptor comprising 2 or more different cytoplasmic signalling

PT components.

XX

PS Disclosure; Fig 4; 90pp; English.

XX

XX This protein comprises a chimeric receptor consisting of an scFv

CC engineered from anti-CD3 human antibody CTMO1 linked to an extracellular

CC spacer in the form of part of human CD8 hinge, linked to the

CC extracellular, transmembrane and intracellular components of the human T

CC cell receptor zeta chain. It can be expressed in host cells (e.g. Jurkat)

CC using a chimeric receptor gene (see AAT90509) constructed from DNA

CC cassettes of each component. In a claimed cell activation process an

CC effector cell is transformed with DNA encoding a chimeric receptor

CC containing 2 or more different cytoplasmic signalling components. Also

CC claimed is use of DNA encoding a recombinant chimeric receptor in a DNA

CC delivery system. The DNA delivery systems can be used for the activation

CC of cells to provide e.g. an increase in cell proliferation, expression of

CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation of

CC cytolytic activity, differentiation or other effector functions, antibody

CC secretion, phagocytosis, tumour infiltration and/or increased adhesion.

CC They can be used in the treatment of e.g. infectious disease,

CC inflammatory disease, cancer, allergic/atopic disease, congenital

CC disease, dermatologic disease, neurologic disease, transplants and

CC metabolic/idiopathic disease (claimed). In particular, they can be used

CC in the treatment of rheumatoid arthritis, osteoarthritis, inflammatory

CC bowel disease, asthma, eczema, cystic fibrosis, sickle cell anaemia,

CC psoriasis, multiple sclerosis, organ or tissue transplant rejection,

CC graft-versus-host disease or diabetes (claimed). (Updated on 17-OCT-2003

CC to standardise OS field)

XX

SQ Sequence 473 AA;

Query Match 67.7%; Score 1590.5; DB 2; Length 473;

Best Local Similarity 66.0%; Pred. No. 6.2e-102;

Matches 318; Conservative 42; Mismatches 63; Indels 59; Gaps 8;

QY 6 IILFLVATATGVHSDIQLTQSPSLASVGRVTITCKASQDV-----GTSVANYQKPG 60

DB 7 VLGLLLWLTDARCDIQMTQSPSTLSASVGRVTITCRSSKSLHSHGNDTFLYWFQKPG 66

QY 61 KAPKLLIYWTSTRHTGVPSRFGSGSGTDFTTTSSIQPEDIAITYYCOQVSLYR-SFQOG 119

DB 67 KAPKLLMYRMSNLASGVPSRFGSGSGTEFTLTSSIQPDDFATYYCMHLEYFTFGQG 126

QY 120 TKVEIKR-----GGSGSGSGSGSGSEVQLVESGGGVYVQPSRLRLSCSAGF 168

DB 127 TKVEVKETGGSG 186

QY 169 DFTTYMWSVRQAPGKLEWIGEIHDPDSSTINVAPSLKDRPTISRDNAKNTFLQMSLR 228

DB 187 TPTDYIINMRQAPGQGLEWIGWIDPGSGNTKYNEKPKGRATLTVDSTNTAYMELSLR 246

QY 229 PEDTGVVFCA----SLYFGFPWFAYWGQGTPTVTVSS-----A 261

DB 247 SEDTAFYFCAREKTTYIYAMD---YWGQGLTVTVSSASTKGPTNSIMYFSHFVFLPA 303

QY 262 KPTTTPAPRPTPTAPTIASQPLSLRPEAARPAAGAVHTRGLDPAIDPCLCYLDGILFI 321

DB 304 KPTTTPAPRPTPTAPTIASQPLSLRPEA-----QSFGL---LQPKLCYLDGILFI 351

QY 322 YGVILTALFLRVKFSRABPPAYQQGQNLNGLRRREYDVLDRKRGRDPEWGKPR 381

DB 352 YGVILTALFLRVKFSRABPPAYQQGQNLNGLRRREYDVLDRKRGRDPEWGKPR 411

QY 382 RKNPQEGLYNELOKQWAEAYSEIGMKGERRRGHDLGYQLSTATKOTYDALHMQALP 441

DB 412 RKNPQEGLYNELOKQWAEAYSEIGMKGERRRGHDLGYQLSTATKOTYDALHMQALP 471

QY 442 PR 443

DB 472 PR 473

RESULT 4

AAW26647

ID AAW26647 standard; protein; 514 AA.

XX

AC AAW26647;

XX

XX 17-OCT-2003 (revised)

DT 12-FEB-1998 (first entry)

XX

DE Chimeric receptor hCTMO1/CD8/zeta-CD28.

XX

KW Cell activation; chimeric receptor; DNA delivery; CTMO1; scFv; CD8; CD28;

KW T cell receptor zeta chain; cell proliferation; cytokine; inflammation;

KW effector; cell differentiation; antibody secretion; phagocytosis;

KW tumour infiltration; adhesion; infection; cancer; allergy;

KW rheumatoid arthritis; osteoarthritis; asthma; eczema;

KW inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;

KW psoriasis; multiple sclerosis; transplant rejection; diabetes;

KW graft versus host disease; human; therapy.

XX

OS Homo sapiens.

OS Chimeric.

XX

PN WO9723613-A2.

XX

PD 03-JUL-1997.

XX

XX 23-DEC-1996; 96WO-GB003209.

XX

XX 21-DEC-1995; 95GB-00026131.

XX

```

XX PA (CLLT ) CELLTech THERAPEUTICS LTD.
XX PI Bebbington CR, Lawson ADG, Weir ANC, Finney HM;
XX DR WPI; 1997-351052/32.
XX DR N-PSDB; AAT90510.
XX PT New DNA systems for activating cells - comprising DNA coding for a
XX PT chimeric receptor comprising 2 or more different cytoplasmic signalling
XX PT components.
XX PS Disclosure; Fig 5; 90pp; English.
XX CC This protein comprises a chimeric receptor consisting of an scFv
XX CC engineered from anti-CD3 human antibody CTMO1 linked to an extracellular
XX CC spacer in the form of part of human CD8 hinge, linked to the
XX CC extracellular, transmembrane and intracellular components of the human T
XX CC cell receptor zeta chain, fused to the intracellular component of human
XX CC CD28. It can be expressed in host cells (e.g. Jurkat) using a chimeric
XX CC receptor gene (see AAT90510) constructed from DNA cassettes of each
XX CC component of the receptor. In a claimed cell activation process an
XX CC effector cell is transformed with DNA encoding a chimeric receptor
XX CC containing 2 or more different cytoplasmic signalling components. Also
XX CC claimed is use of DNA encoding a recombinant chimeric receptor in a DNA
XX CC delivery system. The DNA delivery systems can be used for the activation
XX CC of cells to provide e.g. an increase in cell proliferation, expression of
XX CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation of
XX CC secretory activity, differentiation or other effector functions, antibody
XX CC secretion, phagocytosis, tumour infiltration and/or increased adhesion.
XX CC They can be used in the treatment of e.g. infectious disease,
XX CC inflammatory disease, cancer, allergic/atopic disease, congenital
XX CC disease, dermatologic disease, neurologic disease, transplants and
XX CC metabolic/idiopathic disease (claimed). In particular, they can be used
XX CC in the treatment of rheumatoid arthritis, osteoarthritis, inflammatory
XX CC bowel disease, asthma, eczema, cystic fibrosis, sickle cell anaemia,
XX CC psoriasis, multiple sclerosis, organ or tissue transplant rejection,
XX CC graft-versus-host disease or diabetes (claimed). (Updated on 17-OCT-2003
XX CC to standardise OS field)
XX SQ Sequence 514 AA;
    Query Match 67.7%; Score 1590.5; DB 2; Length 514;
    Best Local Similarity 66.0%; Pred No. 6.9e-102;
    Matches 318; Conservative 42; Mismatches 63; Indels 59; Gaps 8;
    6 IILFLVATATGVHSDIQTQSPSSLSASVSDRVITITCKASQDV-----GTSVAMYQQRPG 60
    7 VLGLLLMLTDARCDIQMTQSPSTLSASVSDRVITITCRSSKLLSHNSGDTFLYWFQKPG 66
    61 KAPKLLIYWTSTRHTGVPSRSGSGGTDFTTISLQPEDIATYTCQOYSLYR-SFQGG 119
    67 KAPKLLMYRMNLSASGVPSRSGSGSGTEFTLTISLQDPDPATYTCMQLHEYPFTFQGG 126
    120 TKVEIKR-----GGSGGGSGGSGSEVLVESGVVQPGVQGRSLRLSCSASGF 168
    127 TKVEVKRTGGGGGGGGGGGGGGGGGGGGGQQLVQSGAEVKKPGSSVKVCSKASGY 186
    169 DFTTYWMSVVRQAPQGLWETIGEIHDPDSTINYAPSLKDRFTISRDNAKNTFLQMDSLR 228
    187 TFTDYINWVRQAPQGLWETIGWIDPGSGNTKYNEKFKGRATLTVDSTNTAYNELSSLR 246
    229 PBDTCGVYCA-----SLYGFQFWFAWGGQTPTVSS-----A 261
    247 SEDTAFYFCAREKTYTYAMD---YWGQGLTVTVSSASTKGPTNSIMYFSHFVFLPFA 303
    262 KPTTTPAPRPTTAPATIASQPLSLRPEAARPAAGCAVHTGLDPAIDPKLCVLLDGLFI 321
    304 KPTTTPAPRPTTAPATIASQPLSLRPEA-----QSFGL---LDPKLCVLLDGLFI 351
    322 YGVILTALFLRVKFSRSRSEPPAYQQGQNLNGLNLRREYDVLDRKRDRDPENGGKPR 381
    352 YGVILTALFLRVKFSRSRSDAPAYQQGQNLNGLNLRREYDVLDRKRDRDPENGGKPR 411
QY 382 RKNPOEGLYNELQDKMAEAYSEIGMKGERRRGKHDGLYQGLSTATKDTYDALHMQALP 441
Db 412 RKNPOEGLYNELQDKMAEAYSEIGMKGERRRGKHDGLYQGLSTATKDTYDALHMQALP 471
QY 442 PR 443
Db 472 PR 473
RESULT 5
AAW73051
ID AAW73051 standard; protein; 532 AA.
XX AC AAW73051;
XX DT 17-OCT-2003 (revised)
XX DT 02-FEB-1999 (first entry)
XX DE Z33dCH2.z chimeric receptor.
XX KW Chimeric receptor; A33; Z33; colorectal cancer; antigen; tumour;
KW cytolysis; gene therapy; signal transduction; monoclonal antibody;
KW single chain antibody; scAb; humanised antibody; human; Z33dCH2.z.
XX OS Mus sp.
XX OS Chimeric.
XX FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= Sig_peptide
FT /note= "MAB 98-6 V kappa signal peptide"
FT Region 23..264
FT /label= Z33
FT /note= "humanised A33 single chain antibody"
FT Peptide 265..274
FT /note= "human gamma 1 heavy chain hinge"
FT Domain 275..379
FT /label= CH3
FT /note= "human gamma 1 CH3 domain"
FT 380..396
FT /label= M1
FT /note= "IGM M1 exon"
FT Domain 397..419
FT /label= Transmembrane
FT /note= "CD4 transmembrane domain"
FT Region 420..532
FT /label= Zeta
FT /note= "intracellular portion of human zeta"
XX WO9841613-A1.
PD 24-SEP-1998.
PF 13-MAR-1998; 98WO-US003797.
PR 14-MAR-1997; 97US-00815030.
XX (OTTE/) OTTEN G R.
PA (GREE/) GREENBURG G B.
PA (CASE/) CASENTINI-BOROCZ D.
PA (FINE/) FINER M H.
XX Otten GR, Greenburg GB, Casentini-Borocz D, Finer MH;
XX WPI; 1998-521211/44.
XX N-PSDB; AAV58937.
XX New chimeric DNA encoding membrane-bound receptor for non-MHC restricted
XX signal activation - inducing cytotoxic effector cells or cytokine
XX production in presence of tumour cells expressing the A33 antigen, and
XX related vectors.

```

Example 3; Page 53; 90pp; English.

This is the amino acid sequence of a chimeric receptor designated Z33dCH2.z, which comprises humanised anti-colorectal cancer antigen A33 single chain antibody Z33 (see AAW73049), human ICG2 CH2 domain and the human zeta intracellular region. The invention provides novel chimeric proteins characterised by an extracellular domain capable of binding to A33 in a non-MHC restricted manner, a transmembrane domain and a cytoplasmic domain capable of activating a signalling pathway. Binding of A33 to the extracellular domain results in transduction of a signal and activation of a signalling pathway in the cell, such that the cell may be induced to carry out various functions relating to the signalling pathway. The chimeric DNA may be used to modify lymphocytes as well as haematopoietic stem cells as precursors to a number of important cell types. The modified cells can be transplanted (especially as a bone marrow transplant) into a mammal to provide a source of cytotoxic effector cells that can kill cells expressing the tumour-associated A33 antigen and cytokine producing cells. The modified cells are activated in presence of A33+ cells. Expression constructs and retrovirus vectors are provided. (Updated on 17-OCT-2003 to standardise OS field)

```

XX      Sequence 532 AA;
PS      Query Match          61.6%; Score 1447; DB 2; Length 532;
XX      Best Local Similarity 59.0%; Pred. No. 6.3e-92;
CC      Matches 310; Conservative 30; Mismatches 77; Indels 108; Gaps 9
CC
CC      This is the amino acid sequence of a chimeric receptor designated
CC      Z33dCH2.z, which comprises humanised anti-colorectal cancer antigen A33
CC      single chain antibody Z33 (see AAW73049), human ICG2 CH2 domain and the
CC      human zeta intracellular region. The invention provides novel chimeric
CC      proteins characterised by an extracellular domain capable of binding to
CC      A33 in a non-MHC restricted manner, a transmembrane domain and a
CC      cytoplasmic domain capable of activating a signalling pathway. Binding of
CC      A33 to the extracellular domain results in transduction of a signal and
CC      activation of a signalling pathway in the cell, such that the cell may be
CC      induced to carry out various functions relating to the signalling
CC      pathway. The chimeric DNA may be used to modify lymphocytes as well as
CC      haematopoietic stem cells as precursors to a number of important cell
CC      types. The modified cells can be transplanted (especially as a bone
CC      marrow transplant) into a mammal to provide a source of cytotoxic
CC      effector cells that can kill cells expressing the tumour-associated A33
CC      antigen and cytokine producing cells. The modified cells are activated in
CC      presence of A33+ cells. Expression constructs and retrovirus vectors are
CC      provided. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ      Sequence 532 AA;
```

QY	16 GVHSDIQLTQSPSSLASVGDRTVITCKASODVGTSVAWYQQKPKGAPKLIIYTSTRHT	75
DB	19 GARCDIQNTQSPSSLSVSGDRTVITCKASNQRVTVAWYQQKPELAKPTLIYLANSRHT	78
QY	76 GVPFRFSGSGSGTDFTFISSLQPEDIATYYCQQYSLY-RSFQGQTKEVIKRGGSGSG-	132
DB	79 GVPFRFSGSGSGTDFTFISSLQPEDIATFYCQHWSYPLTFGQGTKEVEKGSTSGSKP	138
QY	133 GSGLGSSEVQLVESGVGVQGRSLSCSAGSDFDTTWMSWMVRQAPOKGLEWIIEI	192
DB	139 GSGLGSTKGEVOLLESGLVPGGSLRLSCAASGFASFSTDMSVMVRQAPOKGLEWATI	198
QY	193 HPDSSTINYAPSLEKDRFTISRDNAKNTLFLOWDSLRPEDTGTYFCASYLFGFPW----	248
DB	199 SSGSYTVYLDVSVKGRFTISRDSKNLTILQMNSLAEDSAIYYCA-----PTTVPFPA	252
QY	249 YWGOGTPVTWSSAKPTTP-----APPPP-----	272
DB	253 YWGOGTLTVSSDKTHTCPPCPGPQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPFDIA	312
QY	273 -----TPAPTIAS-----	280
DB	313 VEWESNGOPENNYKTTPVLVDSDGSFFLYSKLTVDKSRWQNQNVFSCSMWMEALHNHYTQ	372
QY	281 QPLSLRPEAAPGAAGAVHTRGDLDPALDPKLCYLLDGILFIYGVILTALF--LRVKFSRS	338
DB	373 KSLSLSPELQLAESCAEAQDGEDMAL-----IVLGGVAGALLFLTGLGIFFCVRVKFSRS	427
QY	339 AEPPAYOQGNOLYNELNLGRREYDVLDKKRRGRDPENGKGRPRKNPOEGLYNELQOKDM	398
DB	428 ADAPAYOQGNOLYNELNLGRREYDVLDKKRRGRDPENGKGRPRKNPOEGLYNELQOKDM	487
QY	399 AEAYSEIGMKGERRRRGKHGHDGLYQGLSLSTATKDITYDALHMQUALPPR	443
DB	488 AEAYSEIGMKERRRGKHGHDGLYQGLSLSTATKDITYDALHMQUALPPR	532

RESULT 6	
AAM73050	
ID	AAM73050 standard; protein; 643 AA.
XX	
AC	AAM73050;
XX	
DD	17-OCT-2003 (revised)
DT	02-FEB-1999 (first entry)


```

FT      /note= "CD4 transmembrane region"
FT      522..633
FT      /note= "zeta chain"
XX      WO200023573-A2.
XX      27-APR-2000.
XX      20-OCT-1999; 99WO-US024484.
XX      20-OCT-1998; 98US-0105014P.
XX      (CITY ) CITY OF HOPE.
XX      Raubitschek A, Jensen MC, Wu AM;
XX      WPI; 2000-339676/29.
XX      N-PSDB; AAA15019.
XX      Genetically engineered CD20-specific redirected T cells useful for
XX      treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+
XX      acute or chronic leukemia, and autoimmune disease.
XX      Claim 10; Page 53-55; 58pp; English.
XX      The present sequence represents a synthetic CD20-specific chimeric
XX      receptor. The specification describes CD-20 specific redirected T cells
XX      which express and bear on the cell surface membrane a CD20-chimeric
XX      receptor comprising an intracellular signalling domain, a transmembrane
XX      domain and an extracellular domain, the extracellular domain comprising a
XX      CD20-specific receptor. The genetically engineered CD20-specific
XX      redirected T cells are useful for treating a CD20+ malignancy, such as
XX      non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a human
XX      patient having previously undergone myeloablative chemotherapy and stem
XX      cell rescue. The genetically engineered CD20-specific redirected T cells
XX      are also useful for abrogating an untoward B cell function, such as
XX      autoimmune disease (lupus or rheumatoid arthritis) in a patient
XX      SQ      Sequence 633 AA;

Query Match      51.0%; Score 1198; DB 3; Length 633;
Beat Local Similarity 43.5%; Pred. NO. 1.4e-74;
Matches 273; Conservative 60; Mismatches 101; Indels 194; Gaps 17;

QY      6 IILFLVATATGVHSDIQLTQSPSSLSASVGRVITCKASQDVGTSVAVYQKPKAPKL 65
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      VLLWVPGSTG--DIVITQSPAILUSAPGSKVMTCRASSV-NYMDWYQKPKGSSPK 65

QY      66 LIYMTSTRHTGVPFRFSGSGSGTDFTFTISSLPQEDIAITYYCOQYSLY-RSFGQGTKEI 124
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      WIYATSNLASGVPAFRFSGSGSGTSYSLTISRVEAEDATYTCQQWSFNPTFGGKTLEI 125

QY      125 K--RGGSGSGSGSGSGSEVOLVESGGVQVQGRSLRLSCASGDFDTTYMGMVROAP 182
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      KGSTSGSGSGSGSGSGSEVOLQOQSGAELVKPGASVRMSCKASGYTFTSYNMHWKQTP 185

QY      193 KGLEWIGEIHDPDSTINYPASLKDRTFISRDNKAKNTLFLQWDSLRPEDTGVYFCA-SLY 241
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      QGLEWIGAIYFGNGDTSYNQKFGKATLTADKSSSTAYMQLSSITSEDSADYYCARSNY 245

QY      242 FG--PPWFAYWGQGFVTVSSAKPTTTPAPR--PPTAPTATASOP-----L 283
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      YGSSWFVFDVWGAGTTVTVSSLDPKSKDKTHTCCPCAPELLGGPSVFLFPKPKDTLMI 305

QY      284 SLRPEA----- 289
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      SRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDW 365

QY      290 -----ARPA-----AGG-----AVHT-----RGL- 303
DB      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      LNKGEYKCKVKNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFY 425

QY      304 --DFALD-----PKLCYLL-----DGILFIY----- 322

```

/note= "CD4 transmembrane region"
 522..633
 /note= "zeta chain"
 WO200023573-A2.
 27-APR-2000.
 20-OCT-1999; 99WO-US024484.
 20-OCT-1998; 98US-0105014P.
 (CITY) CITY OF HOPE.
 Raubitschek A, Jensen MC, Wu AM;
 WPI; 2000-339676/29.
 N-PSDB; AAA15019.
 Genetically engineered CD20-specific redirected T cells useful for
 treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+
 acute or chronic leukemia, and autoimmune disease.
 Claim 10; Page 53-55; 58pp; English.
 The present sequence represents a synthetic CD20-specific chimeric
 receptor. The specification describes CD-20 specific redirected T cells
 which express and bear on the cell surface membrane a CD20-chimeric
 receptor comprising an intracellular signalling domain, a transmembrane
 domain and an extracellular domain, the extracellular domain comprising a
 CD20-specific receptor. The genetically engineered CD20-specific
 redirected T cells are useful for treating a CD20+ malignancy, such as
 non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a human
 patient having previously undergone myeloablative chemotherapy and stem
 cell rescue. The genetically engineered CD20-specific redirected T cells
 are also useful for abrogating an untoward B cell function, such as
 autoimmune disease (lupus or rheumatoid arthritis) in a patient

RESULT 12
 AAW36845
 ID AAW36845 standard; protein; 444 AA.
 AC AAW36845;
 XX 23-MAR-1998 (first entry)
 DT Single chain T-cell receptor specific for hemagglutinin antigen.
 DE Her-2/neu protein; human leukocyte antigen A2.1; HLA;
 KW cytotoxic T lymphocyte; CTL; immune response; tumour-associated antigen;
 KW T-cell receptor; TCR; tumour treatment.
 OS Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 133..147
 FT /note= "peptide linker"
 XX WO9732603-A1.
 XX 12-SEP-1997.
 XX 05-MAR-1997; 97WO-US003611.
 XX 05-MAR-1996; 96US-0012845P.
 XX (SCRI) SCRIPPS RES INST.
 XX Sherman LA, Lustgarten J;
 WPI; 1997-470496/43.
 N-PSDB; AAT97844.
 Nucleic acid encoding variable regions of HLA-restricted non-human T cell
 receptor specific for tumour antigen - used to identify tumour antigens
 and for tumour therapy.
 Disclosure; Fig 3; 34pp; English.
 The present sequence represents the complete amino acid sequence of a
 single chain T-cell receptor (TCR) which contains variable alpha and beta
 specific for the hemagglutinin antigen, linked through a short peptide
 linker and then fused through a CD8 hinge to the Z chain. This construct
 was used as a model to design TCRs which are specific for the human
 leukocyte antigen. H7 is a peptide that is based on the Her-2/neu
 protein, and is able to inhibit binding of an influenza matrix protein,
 and to elicit an immune response in vivo in mice. H7 peptide is a tumour-
 associated antigen, and was used to immunize a transgenic, non-human
 vertebrate (that has been modified to express at least one human
 leukocyte antigen (HLA)), so that the animal produces cytotoxic T-
 lymphocytes (CTL) which displays HLA-restricted T-cell receptor (TCR)
 specificity for the antigen. Nucleic acid encoding variable regions of
 the alpha and beta chains of such TCRs can be PCR amplified from these
 CTLs. The present sequence is used as a model to produce chimeric TCRs

CC	from these PCR products. Cells expressing recombinant TCR are used to
CC	identify antigens associated with a tumour and to treat tumours in
CC	humans. Transgenic mice are a more convenient source of CTL than the
CC	tumour-infiltrating lymphocytes previously used. TCR can be humanised to
CC	reduce side-reactions and short peptide derivatives of TCR are more
CC	economical to produce than TCR itself, particularly when expressed as a
CC	single-chain molecule rather than as a dimer
XX	
SO	Sequence 444 AA;
	Query Match 46.6%; Score 1094; DB 2; Length 444;
	Best Local Similarity 54.7%; Pred. No. 1.5e-67;
	Matches 248; Conservative 35; Mismatches 106; Indels 64; Gaps 12;
Qy	22 QLTQSPSSLSASVGRVITTCASQDVGT-SVAVYQKPGKAPKLLIYWTSTRTHTGVPSR 80
Db	25 VQVQSPASLVLOEGENAELOCSFS-IFTNQVQFYQRPGGRLVSLLY-----NPSG 74
Qy	81 FSGSGSGTDFT-----FTISLQPEDIATYYCQVYS-----LYRSPGQGTKEIIRGG 120
Db	75 TKQSGRLTSTTVIKERRSSLHISSQITDSGTYLCASSNGSGSNAKLTFGKGTKLSSVSGG 134
Qy	129 SSGSGSGSGSGSGLVES-GGGVQVQGRSLRLSCSASGDFTTYWMSVVRQAPGKGLE 187
Db	135 GSGSGSGSGGGGSAAVTQSPRNKAVTGGKVTLSCNQTNHNNMY---WYRQDTGHGLR 191
Qy	188 WIGEIHPDSSITNYAPSLKDRFTISRDNAKN-TLFLQWDSLRPBDTGVYFCASLYFGFPW 246
Db	192 LIHYSYAGST-EKGDIPGYKASRPSQENFSLILELAT--PSQTSVYFCASGETGTNE 247
Qy	247 FAYWGQGTPTVVS-----AKPTTTPAPRPTTAPTIASQPLSRPEAA 290
Db	248 RLFFGHGHTKLSVLTNSIMYFSPVFLPAKPTTTPAPRPTTAPTIASQPLSRPSS 307
Qy	291 RPAAGGAVHTRGLDPALDPKLCYLLDGLIFTYGVILTALFLRVKFSRSPAPYQQGQ 350
Db	308 R-----DPKLCYLLDGLIFTYGVILTALFLRVKFSRSDAPAYQQGQ 351
Qy	351 LYNELNLGRREEYDVLDRGRDPEMGKPKRKNPQGLYNELQDKWAEAYSIGMKGE 410
Db	352 LYNELNLGRREEYDVLDRGRDPEMGKPKRKNPQGLYNELQDKWAEAYSIGMKGE 411
Qy	411 RRRKGHDGLYQGLSTATKTDYDALHMQALPPR 443
Db	412 RRRKGHDGLYQGLSTATKTDYDALHMQALPPR 444
RESULT 13	
AAW24025	
ID	AAW24025 standard; protein; 443 AA.
XX	
AC	AAW24025;
AC	
DT	25-MAR-2003 (revised)
DT	04-MAR-1998 (first entry)
XX	
XX	Single chain antigen hybrid receptor.
XX	
KW	Hybrid receptor; single chain antigen; gene therapy; diagnosis;
KW	signal conduction; receptor; control region.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	Peptide 1..19
FT	/label= leader_peptide
FT	Protein 20..444
FT	/note= "single chain antigen hybrid receptor"
FT	Region 20..139
FT	/label= VH B 1-8
FT	/note= "variable heavy chain region of B1-8 antibody"
FT	Region 140..154

Ft		/label= (Gly_Ser)3
Ft		/note= "linker region"
Ft	Region	155..264
Ft		/label= VL B 1-8
Ft		/note= "variable light chain region of BI-8 antibody"
Ft	Region	265..309
Ft		/label= CD_8-alpha_hinge
Ft	Domain	310..330
Ft		/label= transmembrane_domain
Ft	Domain	331..444
Ft		/label= cytoplasmic_domain
Xx		
En	W09720938-A2.	
Xx		
Pd	12-JUN-1997.	
Xx		
Pf	03-DEC-1996;	96WO-DE002334.
Pf		
Xx	05-DEC-1995;	95DE-01045351.
Pr		
Xx		
Pa	(UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.	
Xx		
Pi	Martelsmann R, Kulmburg P, Rosenthal F;	
Xx		
Dr	WFI; 1997-319784/29.	
Dr	N-PSDB; AAT77137.	
Xx		
Pt	Cells with hybrid receptor having extracellular and intracellular regions	
Pt	of different origins - useful in gene therapy and diagnosis of tumours.	
Xx		
Xx	Example 3; Fig 4; 46pp; German.	
Px		
Cc	This sequence represents a novel single chain antigen hybrid receptor	
Cc	(HR) and contains an extracellular domain specific for the hapten 4-	
Cc	hydroxy-5-iodo-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha	
Cc	region and the transmembrane and signal-conducting intracellular parts of	
Cc	the CD3-zeta molecule. Such hybrid receptors comprise a receptor part	
Cc	localised on the outside of the cell and specific to a particular signal	
Cc	molecule and a receptor part originating from another receptor, localised	
Cc	on the inside of the cell and capable of setting off a signal inside the	
Cc	cell. The cell should also contain at least one other gene construct with	
Cc	a control region which can interact with the signal sent out by the	
Cc	hybrid receptor and thereby control expression of a transgene bound to	
Cc	this control region. Such cells are useful in gene therapy or for	
Cc	diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)	
Xx		
Sq	Sequence 443 AA;	
	Query Match	44.2%; Score 1038; DB 2; Length 443;
	Best Local Similarity	51.2%; Pred. No. 1.2e-63;
	Matches 243; Conservative	45; Mismatches 123; Indels 64; Gaps 19;
Qy	1 MGWSCILFLVATATGVHSDIOLTSPPSSLSASVGDRVTITCKASDVCTGS--VANVYQQK 58	
Db	1 MGWSCIMLFLAAATATGVHSQVLQDGAEL-VKPGASVKLSCKASYTFTSYMMHWVKQR 59	
Qy	59 PGKAPKLLIIYWTSTRTGVPDSRFGSGSGGTDF-----TFIT-----SSIQP 99	
Db	60 PGRG----LEWI-----GRDPNSGGTKYNEKFKSKATLVTDKPSSAYMQLSSLTS 107	
Qy	100 EDIATYYCOOYSLYRS-----FQGKTKEIKRGSGSGSGSGSGSEVOLVESGGGVQ 154	
Db	108 EDSAVYICARYDYGSSYFDYWGQGTVTTVSSGGGSGGGGGGSGAQTQSALTTTS 167	
Qy	155 PGRSLRLSC-SASGFDTFTYMSWVRQAPG---KGLEWIGEIHPDSTINYPASLKDRFT 210	
Db	168 PGETVTLTCRSSTGAVTTSNYANWVEKPDHLFTGL--IG-----GTNNRAPGVPARFS 219	
Qy	211 ISRDNAKNLTFLQMDSLRPEDTGYFCASLYFGFPFWFYMGQGTPTVTVSASKPTTPAPR 270	
Db	220 GSLIGDKAA--LTITGAQTEDEAIYFCA-LWYNHNW--VFEGGTGLTLVLEFPTTKPVLR 274	
Qy	271 PTTPA-PTIASQPLSRPEARPAAGGAVHTRGLDFALDPKLCYLLDGILFIYGVIITAL 329	

```

SQ Sequence 443 AA;
Query Match          44.2%; Score 1038; DB 2; Length 443;
Best Local Similarity 51.2%; Pred No. 1.2e-63;
Matches 243; Conservative 45; Mismatches 123; Indels 64; Gaps 19;

Qy 1 MGWSCILFLAVATATGVHSDIQLTQSPSSLSASVGRDVTITCKASQDVGTGTS--VAMVQQK 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MGWSCIMFLAAATATGVHSGVQLQQSGAEL-VKPGASVKLSCKRAGTYTFTSYNMHWKQR 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 59 PGKAPKLLIYWTSTRHTGVPSRSGSGSGHDF-----TFTI-----SSLPQ 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 PCRGG-----LEWI-----GRDPNSGGTKYNEKFKSKATLTVDKPSSTAYMQLSLTS 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 100 EDIATYYCOOYSLYRS-----PQCGTKVEIKRGGSGSGSGSGSEVQLVESGGGVQ 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 EDSAVYICARYDYGSSYFDYWGQGTGTTVSSGGGGSGGGGGGQAVGTQESALTTS 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 155 PGRSLRLSC-SASGDFDTTWMSWVRQAPG--KGLEWIGEIHDPDSSTINVASLKDREF 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 PGEIVTLTRCSRSGAVTTSNANWVEKDPHLFTGL--IG-----GTTNRAPGVPARFS 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 211 ISRDNAKNTLFLQMDSLRPEDTGVIYFCASLYFGFPWFAYMGQGTPTVTSSAKPTTTPAPR 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 GSLIGDKAA--LTITGAQTEDEAIYFCA-LWYSNHW--VEGGGTGLTVLEEFYTTKPVLR 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 271 PPTPA-PTIASQPLSRPEAARPAAGGAVHTRGLDFALDPKLCYLDGLGILFIYGVILTAL 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 275 TPSPVHPTGTSP--QRPEDCRPR--GSVKGTGLDF--DPKLCYLLDGILFIYGVITAI 328
 QY 330 FLRVKFSASAPPAQQONQNLVYNELNGRREYDVLDRGRDRPEMGK-PRKNPQEG 388
 Db 329 YLRKFSRSATAANLQDPNQLVYNELNGRREYDVLDRGRDRPEMGKQQRNRNPOEG 388
 QY 389 LYNELQDKMAEAYSEIGMKERRRGKHGDLGYQGLSTATKDTYDALHMQALPPR 443
 Db 389 IYNALQDKNMAEAYSEIGTKERRRGKHGDLGYQGLSTATKDTYDALHMQALPPR 443

RESULT 14

AAR85508

ID AAR85508 standard; protein; 461 AA.

XX AAR85508;

DT 16-MAR-1996 (first entry)

DE Leader-scFv(FRP5):lyt-2 hinge;zeta.

KW Single chain antibody; scFv; antibody engineering; antitumour;
 KW tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;
 KW monoclonal antibody; erbB-2; cancer; cell targeting;
 KW adoptive immunotherapy.

OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 40..93
 FT /label= IGH_chain_leader
 FT Domain 94..819
 FT /label= scFv(FRP)5
 FT Peptide 261..322
 FT /label= Lyt-2_hinge
 FT Domain 323..461
 FT /label= Zeta_chain

XX WO9530014-A1.

XX PD 09-NOV-1995.

XX PF 20-APR-1995; 95WO-EP001494.

XX PR 02-MAY-1994; 94EP-00810244.

XX PA (CIBA) CIBA GEIGY AG.

XX PI Groner B, Moritz D;

XX DR WPI; 1995-393085/50.

XX DR N-PSDB; AAT05783.

XX New bifunctional proteins for use in killing tumour cells - contg. a
 PT tumour antigen binding domain, a hinge region and a zeta chain derived
 PT from a T-cell antigen receptor.

XX Example 1; Page 29-31; 46pp; English.

XX A bifunctional protein (AAR85505) consists of a single chain antibody,
 CC scFv(FRP)5, directed against the tumour erbB-2 antigen, a hinge region,
 CC and a functional zeta chain obtd. from a T-cell receptor. The protein is
 CC expressed in host cells, esp. cytotoxic T-lymphocytes, providing them
 CC with a defined tumour cell specificity enabling targeting to defined
 CC tumour cells and MHC-unrestricted and MHC-independent tumour destruction
 CC in vitro or in vivo

XX Sequence 461 AA;

Query Match

Best Local Similarity 43.2%; Score 1015.5; DB 2; Length 461;

Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

QY 16 GVHSDIQLTQSPSLSASVGDVRVITCKASODVGTSAWYQOKPKAPKLIYWTSTRT 75
 Db 149 GGGSDIQLTQSHKFLSTSVGDRVITCKASQDVNVAWYQOKPQSPKLIYSSASRYT 208
 QY 76 GVPFRFGSGSGTDTFTTISLQPEDIATYCYQVSYLR---SFGQGTKEIKRGGSGG 132
 Db 209 GVPFRFTGSGSGPDFTTISVQAEDLAVFCQGH--FRTPFTFGSGTKLEIK----- 259
 QY 133 GSGSGSGSEVLVESGGVQVQPSRLSCASGDFDTTYWMSWVRQAPKGLEWIGE 192
 Db 260 -----ALEISNSV----- 267
 QY 193 HPDSSTINYAPSLKDRFTISRDNAKNTLFLQMDSLRPEDTGVYFCASLYFGFPWFAYWGQ 252
 Db 268 -----MYESS 272
 QY 253 GTPVTSSAKPTTPPAPRPTPA-PTIASQPLSLRPEAARPAAGAVHTRGLDFAIDPKL 311
 Db 273 VVPVLQKVNSTTKPVLRTPSVHPTGTSP--QRPEDCRPR--GSVKGTGLDFAIDPKL 328
 QY 312 CYLLDGILFIYGVITAIYLRKFSRSASAPPAQQONQNLVYNELNGRREYDVLDRGR 371
 Db 329 CYLLDGILFIYGVITAIYLRKFSRSASAPPAQQONQNLVYNELNGRREYDVLDRGR 388
 QY 372 RDPENGGK-PRKNPQSGLYNELQDKMAEAYSEIGMKERRRGKHGDLGYQGLSTATKD 430
 Db 389 RDPENGGKQQRNRNPOEGVYNALQDKMAEAYSEIGTKERRRGKHGDLGYQGLSTATKD 448
 QY 431 TYDALHMQALPPR 443
 Db 449 TYDALHMQTLAPR 461

RESULT 15

AAY44994

ID AAY44994 standard; protein; 523 AA.

XX AAY44994;

XX DT 23-MAY-2000 (first entry)

XX DE HD70scFv-CH1-GM-CSF chain.

XX HD70; single-chain variable fragment; scFv; 17-1A antigen; human; EpCAM;
 KW epithelial cell adhesion molecule; inflammatory cytokine; GM-CSF;
 KW granulocyte/macrophage colony stimulating factor; heteroninibody;
 KW CH1-domain; multifunctional compound; heavy chain constant domain;
 KW immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;
 KW antiproliferative; prevention; treatment; malignant; haematopoietic cell;
 KW lymphoma; leukaemia; solid tumour; carcinoma; melanoma; sarcoma.

XX OS Homo sapiens.

XX PN WO200006605-A2.

XX PD 10-FEB-2000.

XX PF 28-JUL-1999; 99WO-EP005416.

XX PR 28-JUL-1998; 98EP-00114082.

XX PA (MTCR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.

XX PI Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;

XX DR WPI; 2000-195265/17.

XX DR N-PSDB; AAZ50587.

XX New multifunctional compounds useful for preventing and/or treating
 PT malignant cell growth and for detection and diagnosis.

XX Example 10; Fig 55A; 166pp; English.

XX The patent discloses heteroinibodies which are multifunctional compounds
CC producible in a mammalian host cell as a secretable and fully functional
CC heterodimer of two polypeptide chains, where one of the polypeptide
CC chains comprises, a CH1-domain (constant domain of an immunoglobulin
CC heavy chain) and the other chain comprises C1-domain (constant domain of
CC an immunoglobulin light chain). The polypeptide chains further comprise,
CC fused to the constant domains at least two (poly)peptides having
CC different receptor or ligand functions, where further at least two of the
CC different (poly)peptides lack an intrinsic affinity for one another and
CC are linked via the constant domains. The heteroinibodies have
CC cytostatic, immunostimulatory, antileukaemia and antiproliferative
CC activities. These compounds can be used for diagnosing, preventing and
CC treating malignant cell growth related to malignancies of haematopoietic
CC cells e.g. lymphomas and leukaemias, or to solid tumours e.g. carcinomas,
CC melanomas and sarcomas. The present sequence is the left chain of a
CC heteroinibody comprising HD70 single-chain Fv (scFv) fragment N-
CC terminally linked to human CH1 domain which bears at its C-terminus the
CC human inflammatory cytokine granulocyte/macrophage colony stimulating
CC factor (GM-CSF), plus a hexahistidine sequence for ease of purification.
CC HD70 scFv specifically recognises the human epithelial cell adhesion
CC molecule (EPCAM) also called 17-1A antigen
XX
XX Sequence 523 AA;

Query Match 42.2%; Score 991; DB 3; Length 523;
Best Local Similarity 64.7%; Pred. No. 2.5e-60;
Matches 198; Conservative 33; Mismatches 61; Indels 14; Gaps 5;

QY 1 MGHSCILFLVATATGVHSDIQLTQSPSSLSASVGRVTITCKASQDVGTISVAVYQOKPG 60
DB 1 MGHSCILFLVATATGVHSDIQLTQSPSSLSASVGRVTITCKASQDVGTISVAVYQOKPG 60

QY 61 KAPKLIYWTSTRHTGVPSRFGSGGCTDFTTISLQPEDIAITYCOQY-SLYRSGQG 119
DB 61 QPPKLIYWTSTRHTGVPSRFGSGGCTDFTTISLQPEDIAITYCOQY-SLYRSGQG 120

QY 120 TKVEIKRGSGSGSGSGSGSEVQLVSGGGVVPGRSLRLSCSASGDFDTTYMWSVR 179
DB 121 TRLDI-QGGGGSGSGSGSGSEVQLVSGGGVVPGRSLRLSCSASGDFDTTYMWSVR 179

QY 180 QAPGKLEWIGEITHPSSTINYPAPSLKDRFTISRDNAKNTLFLQMSLRPEDTGVYFCA- 238
DB 180 QAPGKLEWAVISYDGSNKYIYADSVKGRFTISRDNSKNTLYLQMSLRRAEDTAVYYCAK 239

QY 239 SLVFGFPWFAY-----WGQGTPTVTVSSAKPTTTPAPRPTTAPTIASQPLSLRPEAR 291
DB 240 DMGWSGWRPYYYGMDVWGQGTITVTVSSGLDTHHTASTKGPSV----FPLAPSSKS 295

QY 292 PAAGGA 297
DB 296 TSGGTA 301

Search completed: May 26, 2005, 02:07:03
Job time : 294 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 00:49:19 ; Search time 68 Seconds
(without alignments)
626.824 Million cell updates/sec

Title: US-10-006-771b-2
Perfect score: 2350
Sequence: 1 MGWSCILFLVATATGVHSD.....LSTATKDTYDALHMQALPPR 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: Piri.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	30.8	163	2 A31768	T-cell receptor ze
2	609	25.9	164	2 A40104	T-cell receptor CD
3	605	25.7	166	2 JC4664	T-cell receptor ze
4	605	25.7	166	2 I46424	T-cell surface gly
5	521.5	22.2	188	2 A45089	T-cell receptor io
6	511	21.7	177	2 S54817	T-cell receptor CD
7	511	21.7	206	2 A35900	T-cell receptor CD
8	500.5	21.3	287	4 PC4402	peLB leader/Ig hea
9	488.5	20.8	119	1 AVMSX4	Ig heavy chain v r
10	486.5	20.7	119	1 AVMSJ5	Ig heavy chain v r
11	482	20.5	139	2 A25912	Ig heavy chain pre
12	481	20.5	130	2 S06817	Ig heavy chain v r
13	480.5	20.4	119	1 AVMS76	Ig heavy chain v r
14	475	20.2	118	2 A47329	Ig heavy chain v r
15	473.5	20.1	136	2 S13791	Ig heavy chain v r
16	465	19.8	117	1 G2MSU1	Ig heavy chain v r
17	457	19.4	121	2 A30560	Ig heavy chain v r
18	445.5	19.0	126	1 G1HUKL	Ig heavy chain v-I
19	445	18.9	117	1 G2MS73	Ig heavy chain v-I
20	443.5	18.9	108	1 K1HULY	Ig kappa chain v-I
21	441.5	18.8	123	2 S40331	Ig kappa chain - h
22	441	18.8	107	2 S36272	Ig lambda chain v
23	439.5	18.7	117	2 S46371	Ig kappa chain v-I
24	438.5	18.7	129	2 S52789	Ig kappa chain v r
25	438	18.6	147	2 I37780	Ig variable region
26	437.5	18.6	131	2 S40352	Ig kappa chain v-J
27	436.5	18.6	117	2 S42466	Ig kappa chain v r
28	436.5	18.6	125	2 S40349	Ig kappa chain v-J
29	435.5	18.5	143	2 S23624	Ig heavy chain v r

30	435	18.5	124	2 S40336	Ig kappa chain v-J
31	434.5	18.5	108	2 I39154	Ig kappa chain (BR
32	434.5	18.5	122	2 S31117	Ig heavy chain - h
33	434.5	18.5	125	2 S40333	Ig kappa chain v-J
34	432	18.4	125	2 S30531	Ig heavy chain v r
35	431	18.3	116	1 HVMS44	Ig heavy chain pre
36	431	18.3	121	2 S19666	Ig heavy chain v r
37	430.5	18.3	110	2 S44118	Ig kappa chain v-J
38	430.5	18.3	127	2 S40367	Ig kappa chain v-I
39	429.5	18.3	108	1 K1HUAU	Ig kappa chain v-I
40	429.5	18.3	128	2 S46372	Ig light chain var
41	429	18.3	139	2 I37781	Ig variable region
42	429	18.3	140	2 S06816	Ig heavy chain pre
43	428.5	18.2	114	2 S46390	Ig heavy chain v r
44	428.5	18.2	117	2 S46376	Ig kappa chain v-J
45	428.5	18.2	122	2 S40370	Ig kappa chain - h

ALIGNMENTS

RESULT 1

A31768
T-cell receptor zeta chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: A31768
R;Weissman, A.M.; Hsu, D.; Orloff, D.G.; Modi, W.S.; Seunanez, H.; O'Brien, S.J.; Klausne
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988
A;Title: Molecular cloning and chromosomal localization of the human T-cell receptor zet
A;Reference number: A31768; MUID:89071765; PMID:2974162
A;Accession: A31768
A;Molecule type: mRNA
A;Residues: 1-163 <WEI>
A;Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042
C;Keywords: phosphoprotein; T-cell receptor; transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 30.8%; Score 723; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.4e-40;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	307	LDPKCYLLDGLFIYGVILTAFLRVKFSRSAPPAQQQNLNGLNLRREYDVL	366
DB	27	LDPKCYLLDGLFIYGVILTAFLRVKFSRSAPPAQQQNLNGLNLRREYDVL	86
QY	367	DKRGDPEMGKPRRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHGDLGYQLST	426
DB	87	DKRGDPEMGKPRRKNPQEGLYNELQDKMAEYSEIGMKGRRRGKHGDLGYQLST	146
QY	427	ATKDTYDALHMQALPPR	443
DB	147	ATKDTYDALHMQALPPR	163

RESULT 2

A40104
T-cell receptor CD3 zeta chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C;Accession: A40104; I55293
R;Weissman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.
Science 239, 1018-1021, 1988
A;Title: Molecular cloning of the zeta chain of the T cell antigen receptor.
A;Reference number: A40104; MUID:88145643; PMID:3278377
A;Accession: A40104
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-164 <WEI>
A;Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PIDN:AAA40171.1; PID:G201132
R;Baniyash, M.; Hsu, V.W.; Seidlin, M.F.; Klausner, R.D.
J. Biol. Chem. 264, 13252-13257, 1989

A:Title: The isolation and characterization of the murine T cell antigen receptor zeta c
A:Reference number: I55293; MUID:89327299; PMID:2787796
A:Accession: I55293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-164 <RES>
A:Cross-references: GB:J04967; NID:g556326; PIDN:AAA50301.1; PID:g556327
C:Genetics:
A:Gene: Tcrz
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein

Query Match 25.9%; Score 609; DB 2; Length 164;
Best Local Similarity 80.7%; Pred. No. 6.5e-33;
Matches 121; Conservative 7; Mismatches 18; Indels 4; Gaps 2;

QY 295 GGAHTRGDLPALDKLCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 354
DB 18 GAAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSAPPAQQGQNLN 74
QY 355 LNLGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDMAEAYSEIGMKGERR 413
DB 75 LNLGRREYDVLDRGRDPEMGKQRRNPQGVNQLQDMAEAYSEIGTKGERR 134
QY 414 KGHDGLYQGLSTATKDTYDALHMOALPPR 443
DB 135 KGHDGLYQGLSTATKDTYDALHMOALPPR 164

RESULT 3
JC4664
T-cell receptor zeta chain - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 05-Nov-1999
A:Accession: JC4664
R:Hagens, G.; Galley, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbela
Gene 169, 165-171, 1996
A:Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chain
A:Reference number: JC4663; MUID:96194796; PMID:8647441
A:Accession: JC4664
A:Molecule type: mRNA
A:Residues: 1-166 <HAG>
A:Cross-references: GB:U25689; NID:g1263011; PIDN:AAC48548.1; PID:g1263012
C:Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to se
lymphokine receptor gene expression.
C:Genetics:
A:Gene: tcr-zeta
C:Keywords: GTP binding; signal transduction; T-cell receptor
F:129-146/Region::GDP/GTP-binding

Query Match 25.7%; Score 605; DB 2; Length 166;
Best Local Similarity 80.7%; Pred. No. 1.2e-32;
Matches 121; Conservative 9; Mismatches 14; Indels 6; Gaps 3;

QY 297 AVHTRGDLPALDKLCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 356
DB 20 AAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSAPPAQQGQNLN 76
QY 357 LGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDMAEAYSEIGMK--GERR 413
DB 77 VGRREYAVLDRGGFDPEMGKQRRNPQGVNQLQDMAEAYSEIGMKSDNQR 136
QY 414 KGHDGLYQGLSTATKDTYDALHMOALPPR 443
DB 137 KGHDGLYQGLSTATKDTYDALHMOALPPR 166

RESULT 4
I46424
T-cell surface glycoprotein CD3 zeta chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
A:Accession: I46424; S22980
R:Hein, W.R.; Tunncliffe, A.

Immunogenetics 37, 279-284, 1993
A:Title: Invariant components of the sheep T-cell antigen receptor: cloning of the CD3 e
A:Reference number: I46424; MUID:93131305; PMID:8420837
A:Accession: I46424
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-166 <HEI>
A:Cross-references: UNIPROT:P29329; EMBL:Z12968; NID:g1399; PIDN:CAA78312.1; PID:g1400
C:Keywords: glycoprotein

Query Match 25.7%; Score 605; DB 2; Length 166;
Best Local Similarity 80.0%; Pred. No. 1.2e-32;
Matches 120; Conservative 11; Mismatches 13; Indels 6; Gaps 3;

QY 297 AVHTRGDLPALDKLCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 356
DB 20 AAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSAPPAQQGQNLN 76
QY 357 LGRREYDVLDRGRDPEMGK-PRKNPQGLYNELOKDMAEAYSEIGMK--GERR 413
DB 77 VGRREYAVLDRGGFDPEMGKQRRNPQGVNQLQDMAEAYSEIGMKSDNQR 136
QY 414 KGHDGLYQGLSTATKDTYDALHMOALPPR 443
DB 137 KGHDGLYQGLSTATKDTYDALHMOALPPR 166

RESULT 5
A45089
T-cell receptor CD3 theta chain, alternate splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
A:Accession: A45089; I49587
R:Clayton, L.K.; Diener, A.C.; Lerner, A.; Tse, A.G.; Koyasu, S.; Reinherz, E.L.
J. Biol. Chem. 267, 26023-26030, 1992
A:Title: Differential regulation of T-cell receptor processing and surface expression af
A:Reference number: A45089; MUID:93100325; PMID:1464613
A:Accession: A45089
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-188 <CIA>
A:Cross-references: GB:S51932; NID:g261998; PIDN:AAB24559.1; PID:g261999
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIP:120865)
R:Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.
Eur. J. Immunol. 22, 2135-2140, 1992
A:Title: Human genomic sequences corresponding to murine CD3eta-related transcripts: lac
A:Reference number: I49587; MUID:92347411; PMID:1322304
A:Accession: I49587
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-188 <RES>
A:Cross-references: GB:L03353; NID:g192508; PIDN:AAA37401.1; PID:g192509
C:Keywords: T-cell receptor

Query Match 22.2%; Score 521.5; DB 2; Length 188;
Best Local Similarity 79.7%; Pred. No. 3.8e-27;
Matches 102; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

QY 295 GGAHTRGDLPALDKLCYLLDGLIFLYGVLTALFLRVKFSRSAPPAQQGQNLN 354
DB 18 GAAQSGFL---LDPKLCYLLDGLIFLYGVLTALFLRAKFSRSAPPAQQGQNLN 74
QY 355 LNLGRREYDVLDRGRDPEMGKQRRNPQGVNQLQDMAEAYSEIGMKGERR 414
DB 75 LNLGRREYDVLDRGRDPEMGKQRRNPQGVNQLQDMAEAYSEIGTKGERR 134
QY 415 KGHDGLYQ 422
DB 135 KGHDGLYQ 142

RESULT 6

S54817
T-cell receptor iota precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C:Accession: S54817
R:Nocentini, G.; Ronchetti, S.; Bartoli, A.; Testa, G.; d'Adamo, F.; Riccardi, C.; Migliorini, P.
Submitted to the EMBL Data Library, January 1995
A:Description: T cell receptor iota: an alternatively spliced product of the T cell receptor gene
A:Reference number: S54817
A:Accession: S54817
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-177 <NOC>
A:Cross-references: EMBL:X84237; NID:g809045; PIDN:CAA59015.1; PID:g809046
C:Keywords: T-cell receptor

Query Match 21.7% Score 511; DB 2; Length 177;
Best Local Similarity 79.1%; Pred. No. 1.7e-26;
Matches 102; Conservative 7; Mismatches 16; Indels 4; Gaps 2;

Qy 295 GGAVHTRGLDFALDPKLCYLLDGLIFTYGIVLTALFLRVKFSRSABPPAYQQGQNOLYNE 354
Db 18 GAQAQSFG---LDPKLCYLLDGLIFTYGIVLTALFLRVKFSRSAAANLPDNPOLYNE 74

Qy 355 LNLGRREYDVLRGRDRPEMGCK-PRRKNPQEGLYNELQDKMAEYSEIGMKGERRR 413
Db 75 LNLGRREYDVLEKGRARDPEMGKKOQRNRNPQGVYNALQDKMAEYSEIGTKGERRR 134

Qy 414 GKGHDLGYO 422
Db 135 GKGHDLGYO 143

RESULT 7
A35900
T-cell receptor CD3 eta chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 09-Jul-2004
C:Accession: A35900; A60374; G46522
R:Jin, Y.-J.; Clayton, L.K.; Howard, F.D.; Koyasu, S.; Sieh, M.; Steinbrich, R.; Tarr, G.; Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990
A>Title: Molecular cloning of the CD3eta subunit identifies a CD3zeta-related product in mice
A:Reference number: A35900; MUID:90239005; PMID:2139725
A:Accession: A35900
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-206 <JIN>
A:Cross-references: UNIPROT:P29020; GB:M33158; NID:g1924488; PIDN:AAA37398.1; PID:g309159
R:Ohno, H.; Saito, T.
Int. Immunol. 2, 1117-1119, 1990
A>Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.
A:Reference number: A60374; MUID:91190781; PMID:2150596
A:Accession: A60374
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 144-206 <OHNO>
R:Jensen, J.P.; Cenciarelli, C.; Hou, D.; Rellahan, B.L.; Dean, M.; Weissman, A.M.
J. Immunol. 150, 122-130, 1993
A>Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross reactivity with alpha beta gamma delta epsilon zeta eta subunits
A:Reference number: A46522; MUID:93107707; PMID:8417118
A:Contents: annotation
C:Comment: The functional significance of this alternatively spliced product of the CD3 gene, differs widely among various mammalian species in sequence, length, and even amino acids, differs widely among various mammalian species in sequence, length, and even amino acids; alternative splicing; T-cell receptor; transmembrane protein

Query Match 21.7% Score 511; DB 2; Length 206;
Best Local Similarity 79.1%; Pred. No. 2e-26;
Matches 102; Conservative 7; Mismatches 16; Indels 4; Gaps 2;

Qy 295 GGAVHTRGLDFALDPKLCYLLDGLIFTYGIVLTALFLRVKFSRSABPPAYQQGQNOLYNE 354
Db 18 GAQAQSFG---LDPKLCYLLDGLIFTYGIVLTALFLRVKFSRSAAANLPDNPOLYNE 74

Query Match 21.3%; Score 500.5; DB 4; Length 287;
Best Local Similarity 40.9%; Pred. No. 1.4e-25;
Matches 112; Conservative 45; Mismatches 94; Indels 23; Gaps 7;

Cross-species alignment:

355 INLGRREYDVLDKRRGDDPMGGK-PRKNPQEGLYNELQDKMAEAYSEIGMKERRR 413
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
75 INLGRREYDVLEKRRADPMGGKQRRRNPEGVYNALQDKMAEAYSEIGTKERRR 134
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

414 GKGDGLGY 422
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
135 GKGDGLGY 143

RESULT 8
PC4402
peLB leader/Ig heavy chain anti-NP/linker type 205/alkaline phosphatase fusion protein -
C/Species: synthetic
C/Date: 06-Nov-1998 #sequence_revision 06-Nov-1998 #text_change 06-Nov-1998
C/Accession: PC4402
R/Suzuki, C.; Ueda, H.; Suzuki, E.; Nagamune, T.
J. Biochem. 122, 322-329, 1997
A>Title: Construction, bacterial expression, and characterization of haptan-specific sin-
A/Reference number: PC4402
A/Accession: PC4402
A/Molecule type: DNA
A/Residues: 1-287 <SUZ>
C/Keywords: fusion protein

Query Match 21.3%; Score 500.5; DB 4; Length 287;
Best Local Similarity 40.9%; Pred. No. 1.4e-25;
Matches 112; Conservative 45; Mismatches 94; Indels 23; Gaps 7;

Cross-species alignment:

8 LFLVATATGVHSDIQ--LTQSPSSLSASGVRDTITCKASQDVGIS--VANVQQKEGKA 62
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
12 LLLAAQAPAMADIQA VVTQB-SALTTPSGFTVTLTCSRSTGAIVTSNANWVOERPDHL 70
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

63 PKLLIYWTSRHTGVPGRFSGSGGTDFTTTISQLQPEDIAYYCO-QYSLYRSFGQGTK 121
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
71 FTGLIGTNRAPGVPARFSGSLIGDKAALITGAQTDEALYFCALWTSNHWVFGGETK 130
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

122 VEI-----XRGSGSGSGSGSGSEVOLVESGGVGWPGRSLRLSCSAGFD 169
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
131 LTVLSSADDAKKDADDAKKDDAKDKDG--QVQLQPGAEELVKPGASVKLSCKASGYT 187
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

170 FTYWMSWRQAPQKGLEWIGEIIHPDSTINYPASLKORFTISRDNAKNTLFLQMDSLRP 229
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

188 FTSYMHEWIKRQPGRGLEWIGRIDPNMSGTKYNEKFASKATILTVDPKPSSTAYMOLSLTS 247
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

230 EDTGVVFCASL-YGFEPWFAYWGOGTPVTVSSAK 262
Qy |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
Db |::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:

248 EDGAVTYCARDYDYGSSFYDWGOGTTLTVSSGE 281

RESULT 9
AVMSX4
Ig heavy chain V region (X44) - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Apr-1980 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C/Accession: A02077; S06815
R/Rao, D.N.; Rudikoff, S.; Krutzsch, H.; Potter, M.
Proc. Natl. Acad. Sci. U.S.A. 76, 2890-2894, 1979
A>Title: Structural evidence for independent joining region gene in immunoglobulin heavy
terminating regions.
A/Reference number: A93832; MUID:79223895; PMID:111245
A/Accession: A02077
A/Molecule type: protein
A/Residues: 1-119 <RAO>
A/Cross-references: UNIPROT:P01807
A/Note: This chain was isolated from an IgA myeloma protein that binds galactan
R/Miller III, A.; Glasel, J.A.
J. Mol. Biol. 209, 763-778, 1989
A>Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morp-
A/Reference number: S06815; MUID:90064531; PMID:2555519
A/Accession: S06815
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-118 <Mil>

A:Experimental source: clone IIIB
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:22-96/Disulfide bonds: #status predicted

Query Match 20.8%; Score 486.5; DB 1; Length 119;
Best Local Similarity 75.8%; Pred. No. 3.2e-25;
Matches 91; Conservative 15; Mismatches 11; Indels 3; Gaps 2;

QY 142 EVQLVESGGGVQPGSRSLRLSCASGDFPTTYWMSWVRQAPGKGLWEIGHPSSTINY 201
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVKLESGGGLVQPGSLKLSCAASGDFSRFYWMSWVRQAPGKGLWEIGHPSSTINY 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGPPWFAYWGQTPVTYSS 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDITALYCARLHYGYA--AYWGQTLVTYSA 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
AVMSJ5
Ig heavy chain V region (J539) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Apr-1980 #sequence_revision 30-Apr-1980 #text_change 09-Jul-2004
C:Accession: A02080
R: Rao, D.N.; Rudikoff, S.; Krutzsch, H.; Potter, M.
Proc. Natl. Acad. Sci. U.S.A. 76, 2890-2894, 1979
A:Title: Structural evidence for independent joining region gene in immunoglobulin heavy
etermining regions.
A:Reference number: A93832; MUID:79223895; PMID:111245
A:Accession: A02080
A:Molecule type: protein
A:Residues: 1-119 <RAO>
A:CROSS-references: UNIPROT:P01810
A:Comment: This chain was isolated from a myeloma protein that binds galactan.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:22-96/Disulfide bonds: #status predicted

Query Match 20.7%; Score 486.5; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 4.3e-25;
Matches 90; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 142 EVQLVESGGGVQPGSRSLRLSCASGDFPTTYWMSWVRQAPGKGLWEIGHPSSTINY 201
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVKLESGGGLVQPGSLKLSCAASGDFSRFYWMSWVRQAPGKGLWEIGHPSSTINY 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGPPWFAYWGQTPVTYSS 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDITALYCARLHYGYN--AYWGQTLVTYSA 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 11
A25912
Ig heavy chain precursor V region (W3129) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 23-Jul-1999
C:Accession: A25912
R: Borden, P.; Kabat, E.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 2440-2443, 1987
A:Title: Nucleotide sequence of the cDNAs encoding the variable region heavy and light c
etermining regions.
A:Reference number: A94147; MUID:87175689; PMID:2436230
A:Accession: A25912
A:Molecule type: mRNA
A:Residues: 1-139 <BOR>
A:CROSS-references: GB:M15873; NID:g195268; PIDN:AAA38228.1; PID:g195269
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-18/Domain: signal sequence #status predicted <STG>
F:19-139/Product: Ig heavy chain V region W3129 #status predicted <VAR>
F:33-116/Domain: immunoglobulin homology <IMM>

Query Match 20.5%; Score 482; DB 2; Length 139;
Best Local Similarity 72.0%; Pred. No. 1e-24;
Matches 90; Conservative 15; Mismatches 18; Indels 2; Gaps 1;

QY 138 GSGEVQLVESGGGVQPGSRSLRLSCASGDFPTTYWMSWVRQAPGKGLWEIGHPSST 197
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 15 GVQCEVKVIESGGGLVQPGSLKLSCAASGDFSRFYWMSWVRQAPGKGLWEIGHPSST 74
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 198 TINTAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFGPPWFAYWGQTP 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 75 TINTAPSLKDRFTISRDNKNTLYLQMSKVRSEDITALYCARLGGDLHYAMDYWGQTS 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 256 VTVSS 260
|||||
Db 135 VTVSS 139
|||||

RESULT 12
S06817
Ig heavy chain V region (clone 11C7) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1991 #sequence_revision 11-Nov-1994 #text_change 23-Jul-1999
C:Accession: S06817
R: Miller III, A.; Glaeser, J.A.
J. Mol. Biol. 209, 763-778, 1989
A:Title: Comparative sequence and immunochemical analyses of murine monoclonal anti-morp
A:Reference number: S06815; MUID:90064531; PMID:2555519
A:Accession: S06817
A:Molecule type: mRNA
A:Residues: 1-130 <MIL>
A:CROSS-references: EMBL:X17165; NID:g51917; PIDN:CAA35043.1; PID:g930155
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:22-96/Disulfide bonds: #status predicted

Query Match 20.5%; Score 481; DB 2; Length 130;
Best Local Similarity 70.5%; Pred. No. 1.1e-24;
Matches 93; Conservative 18; Mismatches 13; Indels 8; Gaps 4;

QY 142 EVQLVESGGGVQPGSRSLRLSCASGDFPTTYWMSWVRQAPGKGLWEIGHPSSTINY 201
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVNLES GGGLVQPGSLKLSCAASGDFSRFYWMSWVRQAPGKGLWEIGHPSSTINY 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVPFCASL-YFG--FPWFAYWGQTP 255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TPSLKDRFTISRDNKNTLYLQMSKVRPEDTGLYCAWTKSVYNGSHYYFDVWGAGTT 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 256 VTVSSAKPTTTP 267
|||||:|||||:|||||
Db 121 VTVSSAK--TTP 130
|||||:|||||:|||||

RESULT 13
AVMST6
Ig heavy chain V region (T601) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Apr-1980 #sequence_revision 30-Apr-1980 #text_change 09-Jul-2004
C:Accession: A02078
R: Rao, D.N.; Rudikoff, S.; Krutzsch, H.; Potter, M.
Proc. Natl. Acad. Sci. U.S.A. 76, 2890-2894, 1979
A:Title: Structural evidence for independent joining region gene in immunoglobulin heavy
etermining regions.
A:Reference number: A93832; MUID:79223895; PMID:111245
A:Accession: A02078
A:Molecule type: protein
A:Residues: 1-119 <RAO>
A:CROSS-references: UNIPROT:P01808
A:Comment: This chain was isolated from an IgA myeloma protein that binds galactan.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
F:22-96/Disulfide bonds: #status predicted

Query Match 20.4%; Score 480.5; DB 1; Length 119;
Best Local Similarity 75.0%; Pred. No. 1.1e-24;
Matches 90; Conservative 14; Mismatches 13; Indels 3; Gaps 2;

QY 142 EVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 201
DB 1 EVKLLSEGGGLVOPGGSKLUSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 60
QY 202 APSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTVSS 260
DB 61 TPLSKDKFIISRDNAKNTLYLQMSKVRSEDTALYYCARLGYGY--FDVWGAGTTVTSS 118

RESULT 14

A47329
IG heavy chain V region (PRI) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 17-Jul-1994 #text_change 30-May-1997
C;Accession: A47329
R;Brinkmann, U.; Gallo, M.; Brinkmann, E.; Kunwar, S.; Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. 90, 547-551, 1993
A;Title: A recombinant immunotoxin that is active on prostate cancer cells and that is
A;Reference number: A47329; MUID:93133825; PMID:8421689
A;Accession: A47329
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-118 <BRI>
A;Experimental source: Balb/c
A;Note: sequence modified after extraction from NCBI backbone
A;Note: sequence extracted from NCBI backbone (NCBIN:122874)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;16-99/Domain: immunoglobulin homology <IMV>

Query Match 20.2%; Score 475; DB 2; Length 118;
Best Local Similarity 74.8%; Pred. No. 2.4e-24;
Matches 89; Conservative 12; Mismatches 16; Indels 2; Gaps 1;

QY 142 EVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 201
DB 2 DVQLVESGGGLVOPGGSKLUSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 61
QY 202 APSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTVSS 260
DB 62 TPLSKDKFIISRDNAKNTLYLQMSKVRSEDTALYYCARR--GYVMDYWGQGTSTVSS 118

RESULT 15

S13791
IG heavy chain V region (X-24) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Aug-1996
C;Accession: S13791
R;Arepalli, S.R.; Heller, M.; Glaudemans, C.P.J.
Nucleic Acids Res. 18, 7152, 1990
A;Title: Sequence of the V(H) gene for murine IgA X-24.
A;Reference number: S13791; MUID:91088313; PMID:2124679
A;Accession: S13791
A;Molecule type: DNA
A;Residues: 1-136 <ARE>
A;Cross-references: EMBL:X55984
A;Note: the authors translated the codon GGG for residue 15 as Lys
C;Genetics:
A;Introns: 15/1
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;33-116/Domain: immunoglobulin homology <IMV>

Query Match 20.1%; Score 473.5; DB 2; Length 136;
Best Local Similarity 71.8%; Pred. No. 3.5e-24;
Matches 89; Conservative 14; Mismatches 18; Indels 3; Gaps 2;

QY 138 GSGSEVOLVESGGVOPGSRSLRLSCASGDFDTTYMWSVVRQAPGKLEWIGIHPDSS 197
DB 15 GVOCEYKLLSEGGGLVOPGGSLNLSCAASGDFSRYYMWSVVRQAPGKLEWIGIHPDSS 74
QY 198 TINYAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASL-YFGFPWFAYWQGTPTV 256
DB 75 TINYTESLKD KFIISRDNAKNTLYLQMSKVRSEDTALYYCARLGYGY--FDVWGQGTTL 132
QY 257 TVSS 260
DB 133 TVSS 136

Search completed: May 26, 2005, 02:15:06
Job time : 69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2005, 19:58:54 ; Search time 345 Seconds
 (without alignments)
 657.539 Million cell updates/sec

Title: US-10-006-771B-2

Perfect score: 2350

Sequence: 1 MGWSCILFLVATGVHSD.....LSTATKDTYDALHMQALPPR 443

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	712	30.3	163	1 CD3Z_PIG	Q9XSJ9 sus scrofa
2	701.5	29.9	164	1 CD3Z_HUMAN	P20963 homo sapien
3	626.5	26.7	165	1 CD3Z_RABIT	Q9Tuf8 oryctolagus
4	611	26.0	262	2 Q652T1	Q65211 mus musculus
5	609	25.9	164	1 CD3Z_MOUSE	P24161 mus musculus
6	605	25.7	166	1 CD3Z_SHEEP	P29329 ovis aries
7	511	21.7	206	1 CD3H_MOUSE	P29020 mus musculus
8	488.5	20.8	119	1 HV37_MOUSE	P01807 mus musculus
9	486.5	20.7	119	1 HV40_MOUSE	P01810 mus musculus
10	480.5	20.4	119	1 HV38_MOUSE	P01808 mus musculus
11	465.5	19.8	298	2 Q9QYF0	Q9QYF0 synthetic c
12	465	19.8	117	1 HV41_MOUSE	P01811 mus musculus
13	462	19.7	118	1 HV39_MOUSE	P01809 mus musculus
14	445.5	19.0	126	1 HV3K_HUMAN	P01772 homo sapien
15	445	18.9	117	1 HV42_MOUSE	P01812 mus musculus
16	443.5	18.9	108	1 KV1M_HUMAN	P01605 homo sapien
17	443	18.9	322	2 Q6KAV0	Q6KAV0 homo sapien
18	442	18.8	493	2 Q6GMX2	Q6GMX2 homo sapien
19	435.5	18.5	470	2 Q6PJA4	Q6PJA4 homo sapien
20	435	18.5	236	2 Q6GMX9	Q6GMX9 homo sapien
21	431.5	18.4	108	1 KV1B_HUMAN	P08362 homo sapien
22	431	18.3	116	1 HV36_MOUSE	P01806 mus musculus
23	429.5	18.3	108	1 KV1B_HUMAN	P01594 homo sapien
24	429.5	18.3	108	2 Q9UL70	Q9UL70 homo sapien
25	429.5	18.3	478	2 Q6PI81	Q6PI81 homo sapien
26	429	18.3	147	2 Q9Y509	Q9Y509 homo sapien
27	428.5	18.2	613	2 Q8WUK1	Q8WUK1 homo sapien
28	428	18.2	236	2 Q7Z3Y4	Q7Z3Y4 homo sapien
29	427.5	18.2	108	1 KV1A_HUMAN	P01593 homo sapien
30	426.5	18.1	108	2 Q9UL77	Q9UL77 homo sapien
31	426.5	18.1	573	2 Q8WU38	Q8WU38 homo sapien

32	426	18.1	107	1 KV1D_HUMAN	P01596 homo sapien
33	426	18.1	236	2 Q6GMW1	Q6GMW1 homo sapien
34	426	18.1	236	2 Q6GMX8	Q6GMX8 homo sapien
35	426	18.1	236	2 Q6PIT5	Q6PIT5 homo sapien
36	425.5	18.1	108	1 KV1O_HUMAN	P01607 homo sapien
37	424.5	18.1	108	1 KV1P_HUMAN	P01608 homo sapien
38	424.5	18.1	108	1 KV1V_HUMAN	P04430 homo sapien
39	423.5	18.0	108	1 KV1H_HUMAN	P01600 homo sapien
40	423	18.0	107	2 Q96SA9	Q96SA9 homo sapien
41	422	18.0	236	2 Q6PIH7	Q6PIH7 homo sapien
42	421	17.9	487	2 Q6ZVK0	Q6ZVK0 homo sapien
43	419.5	17.9	108	1 KV1K_HUMAN	P01603 homo sapien
44	419.5	17.9	485	2 Q6PDE8	Q6PDE8 mus musculus
45	419.5	17.9	606	2 Q6GMV2	Q6GMV2 homo sapien

ALIGNMENTS

RESULT 1
 CD3Z_PIG
 ID CD3Z_PIG STANDARD; PRT; 163 AA.
 AC Q9XSJ9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3Z;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Minnesota miniature swine;
 RA Jie H.-B., Yim D., Kim Y.B.;
 RT "The molecular cloning of porcine CD3 zeta.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta (By similarity).
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (By similarity).
 CC -I- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -I- SIMILARITY: Contains 3 ITAM domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF153830; AAD34640.1;
 CC InterPro; IPR003110; ITAM.
 CC Pfam; PF02189; ITAM; 3.
 CC SMART; SM00077; ITAM; 3.
 CC Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 CC SIGNAL 1 21 By similarity.
 CC CHAIN 22 163 T-cell surface glycoprotein CD3 zeta
 CC chain.
 CC DOMAIN 22 30 Extracellular (Potential).
 CC TRANSMEM 31 51 Potential.
 CC DOMAIN 52 163 Cytoplasmic (Potential).
 CC DOMAIN 69 89 ITAM 1.
 CC DOMAIN 107 128 ITAM 2.
 CC DOMAIN 138 158 ITAM 3.

FT DISULFID 32 32 Interchain (Potential).
 FT MOD RES 152 152 Phosphotyrosine (By similarity).
 SQ SEQUENCE 163 AA; 18568 MW; 34898620B67167C7 CRC64;

Query Match 30.3%; Score 712; DB 1; Length 163;
 Best Local Similarity 98.5%; Pred. No. 5.2e-43;
 Matches 135; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 307 LDPKLCYLLDGLFLFYGVILTFALFURVFPSSAEPAYQCGNQNLNLGRREYDVL 366
 |||
 DB 27 LDPKLCYLLDGLFLFYGVILTFALFURVFPSSADAPAYQCGNQNLNLGRREYDVL 86
 |||

QY 367 DKRRGRDPEMGKPRKRNQPGELYNELQDKMAEYSEIGMKGRRRGKHDGLYQGLST 426
 |||
 DB 87 DKRRGRDPEMGKPRKRNQPGELYNELQDKMAEYSEIGMKGRRRGKHDGLYQGLST 146
 |||

QY 427 ATKDTYDALHMQALPPR 443
 |||
 DB 147 ATKDTYDALHMQALPPR 163
 |||

RESULT 2

CD3Z_HUMAN . STANDARD; PRT; 164 AA.

AC P20963; Q8TAX4;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3Z; Synonyms=T3Z, TCRZ;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89071765; PubMed=2974162;
 RA Weissman A.M., Hou D., Orloff D.G., Modi W.S., Seunanez H.,
 RA O'Brien S.J., Klausner R.D.;
 RT "Molecular cloning and chromosomal localization of the human T-cell
 RT receptor zeta chain: distinction from the molecular CD3 complex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9709-9713(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP INTERACTION WITH HIV-1 NEF.
 RX PubMed=10224289;
 RA Xu X.-N., Laffert B., Screaton G.R., Kraft M., Wolf D., Kolanus W.,
 RA Mongkolsapay J., McMichael A.J., Baur A.S.;
 RT "Induction of Fas ligand expression by HIV involves the interaction of

Nef with the T cell receptor zeta chain.";
 RL J. Exp. Med. 189:1489-1496(1999).
 RN [4]
 RP INTERACTION WITH SLA.
 RX MEDLINE=93380595; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;
 RA Tang J., Sawasaki S., Chang J.-H., Burakoff S.J.;
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell
 RT receptor signaling.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).
 RN [5]
 RP INTERACTION WITH DOCK2.
 RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;
 RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,
 RA Tanaka S.;
 RT "DOCK2 mediates T cell receptor-induced activation of Rac2 and IL-2
 RT transcription.";
 RL Biochem. Biophys. Res. Commun. 296:716-720(2002).
 RN [6]
 RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.
 RX PubMed=12522270; DOI=10.1073/pnas.2436191100;
 RA Salomon A.R., Ficarro S.B., Brill L.M., Brinker A., Phung Q.T.,
 RA Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;
 RT "Profiling of tyrosine phosphorylation pathways in human cells using
 RT mass spectrometry.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448(2003).
 RN [7]
 RP STRUCTURE BY NMR OF 136-149.
 RX MEDLINE=93201600; PubMed=7680960; DOI=10.1016/0092-8674(93)90405-F;
 RA Waksman G., Shoelson S.E., Pant N., Cowburn D., Kuriyan J.;
 RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2
 RT domain: crystal structures of the complexed and peptide-free forms.";
 RL Cell 72:779-790(1993).
 CC -!- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or
 CC heterodimers with CD3-eta. Interacts with SLA and SLA2. Interacts
 CC with DOCK2. Interacts with HIV-1 Nef protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CD-3-zeta;
 CC IsoId=P20963-1; Sequence=Displayed;
 CC Name=CD-3-eta;
 CC IsoId=P20963-2; Sequence=Not described;
 CC -!- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (By similarity).
 CC -!- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -!- SIMILARITY: Contains 3 ITAM domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J04132; AAA60394.1; -;
 CC EMBL; BC025703; AAH25703.1; -;
 CC PIR; A31768; A31768.
 CC PDB; 1TCE; NMR; B=136-149.
 CC Genew; HGNC:1677; CD3Z.
 CC H-InvDB; HIX0001296; -;
 CC MIM; 186780; -;
 CC GO; GO:0005886; C:plasma membrane; TAS.
 CC GO; GO:0042101; C:l-cell receptor complex; TAS.
 CC GO; GO:0005515; F:protein binding; IPI.
 CC GO; GO:0042803; F:protein homodimerization activity; NAS.
 CC InterPro; IPR003110; ITAM.
 CC Pfam; PF02189; ITAM; 3.

DR SMART; SM00077; ITAM; 3.
 KW 3D-structure; Alternative splicing; Phosphorylation; Receptor; Repeat;
 KW Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 164 T-cell surface glycoprotein CD3 zeta chain.
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 164 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 108 129 ITAM 2.
 FT DOMAIN 139 159 ITAM 3.
 FT DISULFID 32 32 Interchain (Potential).
 FT MOD RES 83 83 Phosphotyrosine.
 FT MOD RES 111 111 Phosphotyrosine.
 FT MOD RES 123 123 Phosphotyrosine.
 FT MOD RES 142 142 Phosphotyrosine.
 FT MOD_RES 153 153 Phosphotyrosine (By similarity).
 FT CONFLICT 60 61 DA -> EP (in Ref. 1).
 FT CONFLICT 101 101 Missing (in Ref. 1).
 SQ SEQUENCE 164 AA; 18696 MW; 9408260374856EE9 CRC64;
 Query Match 29.9%; Score 701.5; DB 1; Length 164;
 Best Local Similarity 97.8%; Pred. No. 2.9e-42;
 Matches 135; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 Qy 307 LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQONQLYNELNLGRREYDVL 366
 Db 27 LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQONQLYNELNLGRREYDVL 86
 Qy 367 DKRRGRDPENGGKP-RRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHDGLYQGLS 425
 Db 87 DKRRGRDPENGGKPQRRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHDGLYQGLS 146
 Qy 426 TATKDTYDALHMQALPPR 443
 Db 147 TATKDTYDALHMQALPPR 164
 RESULT 3
 CD3Z_RABIT STANDARD; PRT; 165 AA.
 AC Q9TUF8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).
 DE Name=CD3Z;
 GN Oryctolagus cuniculus (Rabbit).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/J x Chbb:HM;
 RA Isono T., Nishimura M.;
 RT "Rabbit CD3 zeta.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.
 CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (By similarity).
 CC -!- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -!- SIMILARITY: Contains 3 ITAM domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB035152; BAA86994.1; -
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21 By similarity.
 FT CHAIN 22 165 T-cell surface glycoprotein CD3 zeta chain.
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 165 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 107 128 ITAM 2.
 FT DOMAIN 140 160 ITAM 3.
 FT DISULFID 32 32 Interchain (Potential).
 FT MOD RES 154 154 Phosphotyrosine (By similarity).
 SQ SEQUENCE 165 AA; 18773 MW; 3183136130BAA4F5 CRC64;
 Query Match 26.7%; Score 626.5; DB 1; Length 165;
 Best Local Similarity 83.0%; Pred. No. 6.5e-37;
 Matches 122; Conservative 5; Mismatches 15; Indels 5; Gaps 2;
 Qy 299 HTRGLDFALDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQONQLYNELNLG 358
 Db 22 HIFGL---LDPKLCYLLDGLIYGVILTAFLRVKFSRAEPAYQQQONQLYNELNLG 78
 Qy 359 RREYDVLDRGRDPEMGKPRKKNPQEGLYNELQDKMAEAYSEIGMKGE--RRRGKG 416
 Db 79 RREYDVLDRGRDPEMGKPRKKNPQEGLYNELQDKMAEAYSEIGMKGENQRREGK 138
 Qy 417 HDGLYQGLSTATKDTYDALHMQALPPR 443
 Db 139 HDGLYQGLSAATKDTYDALHMQALPPR 165
 RESULT 4
 Q65Z11 PRELIMINARY; PRT; 262 AA.
 ID Q65Z11;
 AC Q65Z11;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Anti-HIV-1 reverse transcriptase single-chain variable.
 DE Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hybridoma;
 RX MEDLINE=96211469; PubMed=8648670;
 RA Shaheen F., Duan L., Zhu M., Bagasra O., Pomerantz R.J.;
 RT "Targeting human immunodeficiency virus type 1 reverse transcriptase by intracellular expression of single-chain variable fragments to inhibit early stages of the viral life cycle."
 RL J. Virol. 70:3392-3400(1996).
 DR EMBL; U48716; AAB64342.1; -
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; IG; 2.
 DR SMART; SM00409; IG; 2.
 DR SMART; SM00408; IGc2; 2.
 DR SMART; SM00406; IGV; 2.
 DR PROSITE; PS50835; IG_LIKE; 2.


```

QY 414 GKGDGLYQGLSTATKDYDALHMQALPPR 443
Db 137 GKGDGLYQGLSTATKDYDALHMQALPPR 166

RESULT 7
CD3H MOUSE
ID CD3H MOUSE STANDARD; PRT; 206 AA.
AC P29020;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor
DE T3 eta chain).
DE Name=CD3z; Synonyms=CD3h;
OS Mus musculus. (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90239005; PubMed=2139725;
RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323 (1990).
RN (2)
RP SEQUENCE OF 144-206 FROM N.A.
RX MEDLINE=91271358; PubMed=1828994;
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206 (1991).
RN (3)
RP SEQUENCE OF 144-206 FROM N.A.
RX PubMed=2150596;
RA Ohno H., Saito T.;
RT "CD3 zeta and eta chains are produced by alternative splicing from a
RT common gene.";
RL Int. Immunol. 2:1117-1119 (1990).
RN (4)
RP ERRATUM.
RA Ohno H., Saito T.;
RL Int. Immunol. 4:1339-1339 (1992).
CC -!- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -!- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a
CC heterodimeric form with CD3-zeta subunit. CD3-eta homodimer has
CC not been observed.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=CD-3-eta;
CC IsoId=P29020-1; Sequence=Displayed;
CC Name=CD-3-zeta;
CC IsoId=P24161-1; Sequence=External;
CC -!- SIMILARITY: Belongs to the CD3Z/FCERIG family.
CC -!- SIMILARITY: Contains 3 ITAM domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

DR EMBL; M33158; AAA37398.1; -.
DR EMBL; M76711; AAA40403.1; -.
DR PIR; A35900; A35900.
DR MGI; 88334; Cd3z.
DR InterPro; IPR003110; ITAM.
DR Pfam; PF02189; ITAM; 2.
DR SMART; SM00077; ITAM; 2.
KW Alternative splicing; Direct protein sequencing; Receptor; Repeat;
KW Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 206
FT T-cell surface glycoprotein CD3 eta
FT chain.
FT DOMAIN 22 30
FT Extracellular (Potential).
FT TRANSMEM 31 51
FT POTENTIAL.
FT DOMAIN 52 206
FT Cytoplasmic (Potential).
FT DOMAIN 69 89
FT ITAM 1.
FT DOMAIN 108 129
FT ITAM 2.
FT DOMAIN 139 159
FT ITAM 3.
FT DISULFID 32 32
FT Interchain (Potential).
SQ SEQUENCE 206 AA; 23339 MW; 829256A2CF44E444 CRC64;

Query Match 21.7%; Score 511; DB 1; Length 206;
Best Local Similarity 79.1%; Pred. No. 1.5e-28;
Matches 102; Conservative 7; Mismatches 16; Indels 4; Gaps 2;

QY 295 GGAVHTRGLDLPALDKLCYLDGILFIYGVILTAFLRVKFSRAEPPAYOQGNOLYNE 354
Db 18 GAQAQSFGL---LDPKLCYLDGILFIYGVITALLYLRAKFSRAETAANLQDPNOLYNE 74

QY 355 LNLGRREYDVLDRGRDPEMGKG-PRKNPQBLQYNELQDKNAEAYSIGMKGERRR 413
Db 75 LNLGRREYDVLKRAKRDPEMGKGQQRNRPNQGVYNALQDKNAEAYSIGTKGERRR 134

QY 414 GKGDGLYQ 422
Db 135 GKGDGLYQ 143

RESULT 8
HV37 MOUSE
ID HV37 MOUSE STANDARD; PRT; 119 AA.
AC P01807;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V region X44.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in complementarity-
RT determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
CC -!- MISCELLANEOUS: This chain was isolated from an IGA myeloma protein
CC that binds galactan.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A02077; AVMSX4.
DR HSPF; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 117
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 13246 MW; BC34FC8F31CD41B3 CRC64;

```

Query Match 20.8%; Score 488.5; DB 1; Length 119;
 Best Local Similarity 75.8%; Pred. No. 3e-27;
 Matches 91; Conservative 15; Mismatches 11; Indels 3; Gaps 2;

QY 142 EVLVESGGVQVQGRSLRLSCASGFDFTTYWMSVWRQAPGKLEWIGIHPDSSTINY 201
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 1 EVKLESGGGLVQPGGSLKLSKAASGDFDSRYWMSVWRQAPGKLEWIGIHPDSSTINY 60
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDYALYYCARLHYGYA--AYWGQGTTLVTSA 118
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
 HV40 MOUSE
 ID HV40 MOUSE STANDARD; PRT; 119 AA.
 AC P01810;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Ig heavy chain V region J539.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=79223895; PubMed=111245;
 RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
 RT "Structural evidence for independent joining region gene in
 immunoglobulin heavy chains from anti-galactan myeloma proteins and
 its potential role in generating diversity in complementarity-
 determining regions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
 RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=88217852; PubMed=3449853;
 RA Suh S.W., Bhat T.N., Navia M.A., Cohen G.H., Rao D.N., Rudikoff S.,
 RA Davies D.R.;
 RT "The galactan-binding immunoglobulin Fab J539: an X-ray diffraction
 study at 2.6-A resolution.";
 RL Proteins 1:74-80 (1986).
 CC -1- MISCELLANEOUS: This chain was isolated from a myeloma protein that
 binds galactan.
 DR PIR; A02080; AVMSJ5.
 DR PDB; 2FBJ; X-ray; H-
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 KW 3D-structure; Direct protein sequencing; Immunoglobulin V region.
 FT NON_TER 119 119
 FT STRAND 3 7
 FT STRAND 10 12
 FT TURN 14 15
 FT STRAND 18 25
 FT STRAND 29 31
 FT HELIX 34 39
 FT STRAND 41 42
 FT TURN 45 51
 FT STRAND 53 54
 FT TURN 58 60
 FT STRAND 62 67
 FT TURN 68 72
 FT STRAND 78 83
 FT HELIX 88 90
 FT STRAND 92 100
 FT TURN 101 103
 FT STRAND 104 108
 FT STRAND 112 116
 SQ SEQUENCE 119 AA; 13240 MW; 577B4F1DB675C1F1 CRC64;

Query Match 20.7%; Score 486.5; DB 1; Length 119;
 Best Local Similarity 75.0%; Pred. No. 4.2e-27;
 Matches 90; Conservative 15; Mismatches 12; Indels 3; Gaps 2;

QY 142 EVLVESGGVQVQGRSLRLSCASGFDFTTYWMSVWRQAPGKLEWIGIHPDSSTINY 201
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 1 EVKLESGGGLVQPGGSLKLSKAASGDFDSRYWMSVWRQAPGKLEWIGIHPDSSTINY 60
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDYALYYCARLHYGYN--AYWGQGTTLVTSA 118
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
 HV38 MOUSE
 ID HV38 MOUSE STANDARD; PRT; 119 AA.
 AC P01808;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Ig heavy chain V region T601.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=79223895; PubMed=111245;
 RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
 RT "Structural evidence for independent joining region gene in
 immunoglobulin heavy chains from anti-galactan myeloma proteins and
 its potential role in generating diversity in complementarity-
 determining regions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894 (1979).
 CC -1- MISCELLANEOUS: This chain was isolated from an IgA myeloma protein
 that binds galactan.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
 DR PIR; A02078; AVMS76.
 DR HSSP; P01810; 2FBJ.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00406; IGV; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 KW Direct protein sequencing; Immunoglobulin V region.
 FT DOMAIN 1 112
 FT NON_TER 119 119
 FT STRAND 119 119
 SQ SEQUENCE 119 AA; 13169 MW; BC38CC94E5EA00E8 CRC64;

Query Match 20.4%; Score 480.5; DB 1; Length 119;
 Best Local Similarity 75.0%; Pred. No. 1.1e-26;
 Matches 90; Conservative 14; Mismatches 13; Indels 3; Gaps 2;

QY 142 EVLVESGGVQVQGRSLRLSCASGFDFTTYWMSVWRQAPGKLEWIGIHPDSSTINY 201
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 1 EVKLESGGGLVQPGGSLKLSKAASGDFDSRYWMSVWRQAPGKLEWIGIHPDSSTINY 60
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 202 APSLKDRFTISRDNKNTLFLQMSLRPDTGYVFCASL-YFGFPWFAYWGQGTPTVTSS 260
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 61 TPSLKDKFTISRDNKNTLYLQMSKVRSEDYALYYCARLHYGY--FDVWGAGTTVTSS 118
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 11
 Q9QYF0
 ID Q9QYF0 PRELIMINARY; PRT; 298 AA.
 AC Q9QYF0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CN 8 single chain antibody.
 GN Name=CN 8 scFv;
 OS synthetic construct.

```
OC other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2018931; PubMed=10706631; DOI=10.1073/pnas.050582197;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a vascular cell wall-specific monoclonal antibody
RT recognizing a cell polarity by using a phase display subtraction
RT method.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR EMBL; AB036341; BAA88633.1; -.
DR PIR; A33933; A33933.
DR PIR; S19112; S19112.
DR HSSP; P01820; 1A70.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 2.
DR SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 19.8%; Score 465.5; DB 2; Length 298;
Best Local Similarity 71.9%; Pred. No. 4e-25;
Matches 87; Conservative 17; Mismatches 14; Indels 3; Gaps 2;

QY 141 SEVLVESGGGVQPGSRSLRLSCSASGDFFTYMSWVRQAPGKGLWIGEIHPDSSTIN 200
Dy :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 39 AQVKLQSGGGLVQPGGSLKLSCAASGDFSRYSWVRQAPGKGLWIGEIHPDSSTIN 98
QY 201 YAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWFAYWGQGTPTVTS 259
Dy :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 99 YTPSLKDKFIISRDNAKNTLYLQMSKVRSEDITLYFCARASYYGHS--AYWGQGTPTVTS 156
QY 260 S 260
Dy 157 S 157

RESULT 12
HV41 MOUSE
ID HV41_MOUSE STANDARD; PRT; 117 AA.
AC P01811;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig heavy chain V region UP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8302113; PubMed=6181731;
RA Auftray C., Sikorav J.L., Ollo R., Rougeon F.;
RT "Correlation between D region structure and antigen-binding
RT specificity: evidences from the comparison of closely related
RT immunoglobulin VH sequences.";
RL Ann. Immunol. (Paris) 132D:77-88(1981).
CC -!- MISCELLANEOUS: This chain was isolated from an Ig gamma-2a myeloma
CC protein binding 2,6-levan.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M36631; AAA38078.1; -.
DR PIR; A02081; G2MSU1.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.

other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2018931; PubMed=10706631; DOI=10.1073/pnas.050582197;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a vascular cell wall-specific monoclonal antibody
RT recognizing a cell polarity by using a phase display subtraction
RT method.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR EMBL; AB036341; BAA88633.1; -.
DR PIR; A33933; A33933.
DR PIR; S19112; S19112.
DR HSSP; P01820; 1A70.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS0835; IG_LIKE; 2.
DR SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 19.8%; Score 465; DB 1; Length 117;
Best Local Similarity 73.0%; Pred. No. 1.4e-25;
Matches 89; Conservative 11; Mismatches 14; Indels 8; Gaps 2;

QY 142 EVQLVESGGGVQPGSRSLRLSCSASGDFFTYMSWVRQAPGKGLWIGEIHPDSSTINY 201
Dy ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 1 EVKLVESGGGLVQPGGSLKLSCAASGDFSRYSWVRQAPGKGLWIGEIHPDSSTINY 60
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWF--FAYWGQGTPTVTV 258
Dy ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 61 TPFLKDKFIISRDNAKNTLYLQMSKVRSEDITLYFCAR----NWDVGFYWGQVTTLTIV 115
QY 259 SS 260
Dy ||
Dy 116 SS 117

RESULT 13
HV39 MOUSE
ID HV39_MOUSE STANDARD; PRT; 118 AA.
AC P01809;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ig heavy chain V region X24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79223895; PubMed=111245;
RA Rao D.N., Rudikoff S., Krutzsch H., Potter M.;
RT "Structural evidence for independent joining region gene in
RT immunoglobulin heavy chains from anti-galactan myeloma proteins and
RT its potential role in generating diversity in complementarity-
RT determining regions.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:2890-2894(1979).
CC -!- MISCELLANEOUS: This chain was isolated from an IgA myeloma protein
CC that binds galactan.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC PIR; P01394; AVMSX2.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR PIR; P00047; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
DR Direct protein sequencing; Immunoglobulin V region.
KW DOMAIN 1 111
FT NON TER 118 118
SQ SEQUENCE 118 AA; 13105 MW; BB16A2DB677EF17F CRC64;

Query Match 19.7%; Score 462; DB 1; Length 118;
Best Local Similarity 73.3%; Pred. No. 2.3e-25;
Matches 88; Conservative 14; Mismatches 14; Indels 4; Gaps 3;

QY 142 EVQLVESGGGVQPGSRSLRLSCSASGDFFTYMSWVRQAPGKGLWIGEIHPDSSTINY 201
Dy ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 1 EVKLVESGGGLVQPGGSLKLSCAASGDFSRYSWVRQAPGKGLWIGEIHPDSSTINY 59
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWFAYWGQGTPTVTVSS 260
Dy ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Dy 60 TPFLKDKFIISRDNAKNTLYLQMSKVRSEDITLYFCARLYGYG--FDYWGQGTTLTVSS 117
```

```

Query Match 19.0%; Score 445.5; DB 1; Length 126;
Best Local Similarity 71.3%; Pred. No. 3.7e-24;
Matches 92; Conservative 9; Mismatches 15; Indels 13; Gaps 3;

QY 142 EVQLVESGGGVQPGSRSLRSCSASGPDFTTYYMSWVRQAPGKLEWIGHDPDSSTINY 201
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 QVQLVESGGGVQPGSRSLRSCSASGPDFTTYYMSWVRQAPGKLEWIGHDPDSSTINY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 202 APSLKORFTISRDNKNTLFLQMDSLRPEDTGVYFCA-----SLYGFPPFPAYWG 251
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADSVKGRFTISRDNKNTLFLQMDSLRPEDTGVYFCA-----SLYGFPPFPAYWG 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 QGTPVTVSS 260
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 QGTPVTVSS 126

RESULT 15
HV42_MOUSE
ID HV42_MOUSE STANDARD; PRT; 117 AA.
AC P01812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ig heavy chain V region MOPC 173.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-104.
RX MEDLINE=72105531; PubMed=5062012;
RA Bourgois A., Fougereau M., de Preval C.;
RT "Sequence of amino acids of the NH 2 -terminal region of a mouse-
   clonal immunoglobulin heavy chain.";
RL Eur. J. Biochem. 24:446-455 (1972).
RN [2]
RP SEQUENCE OF 105-117.
RX MEDLINE=76091933; PubMed=812695;
RA Rocca-Serra J., Milili M., Fougereau M.;
RT "Determination of the primary structure of a mouse IgG2a
   immunoglobulin. Amino-acid sequence of the H4 cyanogen-bromide
   fragment.";
RL Eur. J. Biochem. 59:511-523 (1975).
RN [3]
RP SEQUENCE OF 96-117 FROM N.A.
RX MEDLINE=81223769; PubMed=6787590;
RA Gough N.W., Bernard O.;
RT "Sequences of the joining region genes for immunoglobulin heavy chains
   and their role in generation of antibody diversity.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:509-513 (1981).
RN [4]
RP DISULFIDE BOND.
RX PubMed=11947590;
RA Bourgois A., Fougereau M.;
RT "Partial amino acid sequence of the variable region of a mouse
   gammaG2a immunoglobulin heavy chain. Evidence for the existence of a
   third sub-group of variability for the heavy chain pool.";
RL FEBS Lett. 8:285-288 (1970).
CC -1- MISCELLANEOUS: This gamma-2a chain was isolated from a myeloma
   protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR; A91190; G2MS73.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Direct protein sequencing; Immunoglobulin V region.
FT DOMAIN 1 116 Ig-like.
FT DISULFID 22 96
FT CONFLICT 105 105 N -> D (in Ref. 2).

```

FT NON TER 117 117
SQ SEQUENCE 117 AA; 13051 MW; 156DCCC259380F19 CRC64;
Query Match 18.9%; Score 445; DB 1; Length 117;
Best Local Similarity 70.2%; Pred. No. 3.7e-24;
Matches 85; Conservative 16; Mismatches 14; Indels 6; Gaps 2;
QY 142 EVQLVESGGGVVQPGSRSLRLSCASAGDFTTYMSWVRQAPGKGLEWIGEIHPDSSTINY 201
Db 1 EVKLLESGGPLVQLGGSLKLSCAASGDFDSRYWMSWVRQAPGKGLEWIGEIDPNSSTINY 60
QY 202 APSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWFA--YWGQGTPTVTVS 259
Db 61 TPSLKDKFTISRDNKNTLFLQMSKVRSEDALYYCAR----SPYYAMNYWGQGTPTVTVS 116
QY 260 S 260
Db 117 S 117

Search completed: May 26, 2005, 02:12:53
Job time : 347 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 01:18:54 ; Search time 95 Seconds
(without alignments)
348.100 Million cell updates/sec

Title: US-10-006-771B-2

Perfect score: 2350

Sequence: 1 MWSCILFLVATATGVHSD.....LSTAKDYDALHMQALPPR 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfilesl.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	965	41.1	352	4	US-09-203-958A-2
2	871	37.1	283	3	US-09-420-592A-6
3	871	37.1	283	4	US-09-985-442-6
4	871	37.1	283	4	US-09-983-580-6
5	870	37.0	263	3	US-09-069-821-3
6	870	37.0	263	4	US-09-956-086-3
7	870	37.0	263	4	US-09-956-087-3
8	838	35.7	332	4	US-09-135-121B-7
9	824	35.1	304	3	US-08-862-124-14
10	819.5	34.9	240	1	US-08-488-113B-147
11	819.5	34.9	240	1	US-08-477-484B-147
12	819.5	34.9	240	2	US-08-646-360-147
13	819.5	34.9	240	3	US-08-939-765-147
14	819.5	34.9	240	3	US-09-136-389-147
15	819.5	34.9	240	3	US-09-610-838-147
16	819.5	34.9	240	4	US-09-711-485-147
17	808	34.4	491	4	US-10-011-125A-2
18	795	33.8	553	2	US-08-661-052-16
19	795	33.8	553	3	US-09-188-082-16
20	795	33.8	553	3	US-09-364-088-16
21	795	33.8	553	3	US-09-102-716-16
22	787	33.5	354	4	US-09-393-627B-28
23	774.5	33.0	374	4	US-09-646-028-15
24	772	32.9	456	4	US-09-495-880A-11
25	769.5	32.7	265	3	US-09-420-592A-5
26	769.5	32.7	265	4	US-09-985-442-5
27	769.5	32.7	265	4	US-09-983-580-5

28	767	32.6	599	1	US-08-442-542-18	Sequence 18, Appl
29	767	32.6	599	3	US-08-765-469-18	Sequence 18, Appl
30	766.5	32.6	245	3	US-09-069-821-5	Sequence 5, Appl
31	766.5	32.6	245	4	US-09-956-086-5	Sequence 5, Appl
32	766.5	32.6	245	4	US-09-956-087-5	Sequence 5, Appl
33	765.5	32.6	241	2	US-08-224-591-18	Sequence 18, Appl
34	765.5	32.6	241	2	US-08-926-789-18	Sequence 18, Appl
35	763.5	32.5	287	3	US-08-862-124-17	Sequence 17, Appl
36	750.5	31.9	237	2	US-08-224-591-16	Sequence 16, Appl
37	750.5	31.9	237	2	US-08-926-789-16	Sequence 16, Appl
38	748	31.8	532	2	US-08-417-495-6	Sequence 6, Appl
39	748	31.8	532	2	US-08-284-391B-6	Sequence 6, Appl
40	748	31.8	532	3	US-09-218-950-6	Sequence 6, Appl
41	748	31.8	532	4	US-08-394-388A-6	Sequence 6, Appl
42	748	31.8	532	5	PCT-US92-01785-6	Sequence 6, Appl
43	748	31.8	532	5	PCT-US95-00454-6	Sequence 6, Appl
44	740	31.5	239	2	US-08-860-174A-2	Sequence 2, Appl
45	737	31.4	244	2	US-08-392-338A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-203-958A-2

; Sequence 2, Application US/09203958A

; Patent No. 6682928

; GENERAL INFORMATION:

; APPLICANT: KELLER, Tibor

; APPLICANT: GOLDSTEIN, Joel

; APPLICANT: GRAZIANO, Robert

; APPLICANT: DEO, Yashwant M.

; TITLE OF INVENTION: CELLS EXPRESSING ANTI-FC RECEPTOR

; TITLE OF INVENTION: BINDING COMPONENTS

; FILE REFERENCE: MXI-099CPA

; CURRENT APPLICATION NUMBER: US/09/203,958A

; CURRENT FILING DATE: 1998-12-02

; PRIOR APPLICATION NUMBER: 60/067232

; PRIOR FILING DATE: 1997-12-02

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic construct

US-09-203-958A-2

Query Match 41.1%; Score 965; DB 4; Length 352;
Best Local Similarity 75.3%; Pred. No. 2.3e-68;
Matches 195; Conservative 13; Mismatches 43; Indels 8; Gaps 3;

QY	12	ATATGVHSDIQLTQSPSSLSASVGDRTVTTCASQDVGTIS-----VANTYQKPGKAPKL	65
DB	30	ACAQPARSDIQLTQSPSSLSASVGDRTVTTCSSQSVLYSSNQKYLAWYQKPGKAPKL	89
QY	66	LIYWTSTRHTGVPSRFGSGSGTDFTTTSSLPEDIAITYCOQSYLYRSFGGQTKVEIK	125
DB	90	LIYWASTRSGVPSRFGSGSGTDFTTTSSLPEDIAITYCHQLSSWTFGGQTKVEIK	149
QY	126	-RGSGSGSGSGSGSGSEVQLVESGGGVVQPGKSLRSLSCSASGDFDTTWMVVRQAPGK	184
DB	150	SSGSGSGSGSGSGSGSEVQLVESGGGVVQPGKSLRSLSCSASGDFDTTWMVVRQAPGK	209
QY	185	GLEWGEIHPPDSTINAPSLKDRFTISRDNAKNTLFLQWDSLRPDTGYVFCASLYFGP	244
DB	210	GLEWVATISDGGSYTYTPDSVKGRTISRDNSKNTLFLQWDSLRPDTGYVFCARGYIRY	269
QY	245	PN-FAYWGQGTPTVYSSAK	262
DB	270	EGANDYWGQGTPTVYSSPR	288

```
RESULT 2
US-09-420-592A-6
; Sequence 6, Application US/09420592A
; Patent No. 6333396
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6333396el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300001
; CURRENT APPLICATION NUMBER: US/09/420.592A
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
; US-09-420-592A-6
Query Match 37.1%; Score 871; DB 3; Length 283;
Best Local Similarity 68.2%; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;
QY 20 DIQTQSPSSLSASVGDRTTTCRASQ---DVGTSVAVYQKPGKAPKLLIYWTSTRTHTG 76
Db 1 DIQTQSPSSLSASVGDRTTTCRASQSLVSIISNYLAWYQKPGKAPKLLIYAASLSG 60
QY 77 VPSRFGSGSGTDFTFITSSLOPEDATYYCOQYSLYR--SFGQGTKEIKRGSGSG-- 132
Db 61 VPSRFGSGSGTDFTFITSSLOPEDATYYCOQYNSLPFWTFGQGTKEIKRGSGSGKP 120
QY 133 GSGSGSGSEVOLVESGGVQPGSRSLRSCASGDFDTTYMWSVWROAPKGLWIGEI 192
Db 121 GSGSGSTKGEVOLVESGGVQPGSLRLSCAASGFTFSSVAMSWVROAPKGLWISVI 180
QY 193 --HPDSSTINYAPSLKDRFTISRDNKNTLFLQWMSLRPDTGVYFCA-----SLY 241
Db 181 SKGTDGGSYYVADSVKGRFTISRDNKNTLYLQWNSLRRAEDTAVVYCARGKXGSLSGY 240
QY 242 FGFPWFAYWGQGTPTVYSSAK 262
Db 241 YYHYFDYWGQGTTLVTSSKK 261
RESULT 3
US-09-985-442-6
; Sequence 6, Application US/09985442
; Patent No. 6692942
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6692942el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300003
; CURRENT APPLICATION NUMBER: US/09/985.442
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/420,592
```

```
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
; US-09-985-442-6
Query Match 37.1%; Score 871; DB 4; Length 283;
Best Local Similarity 68.2%; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;
QY 20 DIQTQSPSSLSASVGDRTTTCRASQ---DVGTSVAVYQKPGKAPKLLIYWTSTRTHTG 76
Db 1 DIQTQSPSSLSASVGDRTTTCRASQSLVSIISNYLAWYQKPGKAPKLLIYAASLSG 60
QY 77 VPSRFGSGSGTDFTFITSSLOPEDATYYCOQYSLYR--SFGQGTKEIKRGSGSG-- 132
Db 61 VPSRFGSGSGTDFTFITSSLOPEDATYYCOQYNSLPFWTFGQGTKEIKRGSGSGKP 120
QY 133 GSGSGSGSEVOLVESGGVQPGSRSLRSCASGDFDTTYMWSVWROAPKGLWIGEI 192
Db 121 GSGSGSTKGEVOLVESGGVQPGSLRLSCAASGFTFSSVAMSWVROAPKGLWISVI 180
QY 193 --HPDSSTINYAPSLKDRFTISRDNKNTLFLQWMSLRPDTGVYFCA-----SLY 241
Db 181 SKGTDGGSYYVADSVKGRFTISRDNKNTLYLQWNSLRRAEDTAVVYCARGKXGSLSGY 240
QY 242 FGFPWFAYWGQGTPTVYSSAK 262
Db 241 YYHYFDYWGQGTTLVTSSKK 261
RESULT 4
US-09-983-580-6
; Sequence 6, Application US/09983580
; Patent No. 6764853
; GENERAL INFORMATION:
; APPLICANT: Filpula, David R.
; APPLICANT: Wang, Maoliang
; APPLICANT: Whitlow, Marc D.
; TITLE OF INVENTION: No. 6764853el Method for Targeted Delivery of Nucleic Acids
; FILE REFERENCE: 0977.2300002
; CURRENT APPLICATION NUMBER: US/09/983.580
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 09/420,592
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,949
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Kabat
; OTHER INFORMATION: Consensus
```

```

; NAME/KEY: UNSURE
; LOCATION: (232)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (234)
; OTHER INFORMATION: May be any amino acid.
; NAME/KEY: UNSURE
; LOCATION: (239)
; OTHER INFORMATION: May be any amino acid.
US-09-983-580-6

Query Match      37.1%; Score 871; DB 4; Length 283;
Best Local Similarity 68.2%; Pred. No. 4.7e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITTCASQ---DVGTSAVWYQKPKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGRVITTCASQSLVSIISNYLAWYQKPKAPKLLIYAASLSLG 60

QY 77 VPSRFGSGSGTDTFTTISLQPEDFATYYCQOYSLYR--SFGQGTKEIKRGGSGG-- 132
Db 61 VPSRFGSGSGTDTFTTISLQPEDFATYYCQOYSLPWTFGQTKVEIKGSTSGSGKP 120

QY 133 GSGSGGSGEVQVLESQVQGRSLRLSCSAGDFFTTYMWSWVRQAPKGLWIGEI 192
Db 121 GSGEGSTKGEVQVLESQVQGRSLRLSCAASGFTFSSYAMSWVRQAPKGLWYSVI 180

QY 193 --HPDSSTINAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCA-----SLY 241
Db 181 SGKTGGSTYADSVKGRFTISRDNKNTLYLQMSLRABDTAVYYCARGXKXSLSGXY 240

QY 242 FGPFWFAYWQGTPTVTSSAK 262
Db 241 YYHYFDYWGQGLTVTVSSKK 261

RESULT 5
US-09-069-821-3
; Sequence 3, Application US/09069821
; Patent No. 632322
; GENERAL INFORMATION:
; APPLICANT: FILPULA, DAVID
; APPLICANT: WANG, MAOLIANG
; APPLICANT: SHORR, ROBERT
; APPLICANT: WHITLOW, MARC
; APPLICANT: LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/069,821
; FILING DATE: 30-APR-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/067,341
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/063,074
; FILING DATE: 27-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/050,472

```

```

; FILING DATE: 23-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/044,449
; FILING DATE: 30-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KIM, JUDITH U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 0977.2280003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)371-2600
; TELEFAX: (202)371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-069-821-3

Query Match      37.0%; Score 870; DB 3; Length 263;
Best Local Similarity 68.2%; Pred. No. 5.2e-61;
Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITTCASQ---DVGTSAVWYQKPKAPKLLIYWTSTRHTG 76
Db 1 DIQTQSPSSLSASVGRVITTCASQSLVSIISNYLAWYQKPKAPKLLIYAASLSLG 60

QY 77 VPSRFGSGSGTDTFTTISLQPEDFATYYCQOYSLYR--SFGQGTKEIKRGGSGG-- 132
Db 61 VPSRFGSGSGTDTFTTISLQPEDFATYYCQOYSLPWTFGQTKVEIKGSTSGSGKP 120

QY 133 GSGSGGSGEVQVLESQVQGRSLRLSCSAGDFFTTYMWSWVRQAPKGLWIGEI 192
Db 121 GSGEGSTKGEVQVLESQVQGRSLRLSCAASGFTFSSYAMSWVRQAPKGLWYSVI 180

QY 193 --HPDSSTINAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCA-----SLY 241
Db 181 SGKTGGSTYADSVKGRFTISRDNKNTLYLQMSLRABDTAVYYCARGXKXSLSGXY 240

QY 242 FGPFWFAYWQGTPTVTSSAK 262
Db 241 YYHYFDYWGQGLTVTVSSNK 261

RESULT 6
US-09-956-086-3
; Sequence 3, Application US/09956086
; Patent No. 6743896
; GENERAL INFORMATION:
; APPLICANT: FILPULA, DAVID
; APPLICANT: WANG, MAOLIANG
; APPLICANT: SHORR, ROBERT
; APPLICANT: WHITLOW, MARC
; APPLICANT: LEE, LIHSYNG S.
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/956,086
; FILING DATE: 20-Sep-2001

```

CLASSIFICATION: <Unknown>

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 09/069,821
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: No. 6743896 Relevant
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-956-086-3

Query Match 37.0%; Score 870; DB 4; Length 263;

Best Local Similarity 68.2%; Pred. No. 5.2e-61;

Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITCKASQ---DVGTSVAMYQKPKGKAPKLLIYWTSTRHTG 76

Db 1 DIQTQSPSSLSASVGRVITCKASQSLVSIINLAWYQKPKGKAPKLLIYAASSLESG 60

QY 77 VPSRFGSGSGCTDFTTISLQPEDFATYYCQYSLYR--SFGQGTKEIKRGGSGG-- 132

Db 61 VPSRFGSGSGCTDFTTISLQPEDFATYYCQYNSLPFWTGGQTKVEIKGTSGSGKP 120

QY 133 GSGGSGSGSEVQLVESGGVQVPGKSLRLSCASGFDFTTVMWVNRQAPKGLWIGEI 192

Db 121 GSGEGSTKGEVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSVNRQAPKGLWVSVI 180

QY 193 --HPDSSTINVAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCA-----SLY 241

Db 181 SKGTDGGSTYYADSVKGRFTISRDNKNTLYLQWNSLRAEDTAVYYCARGXGXLSGX 240

QY 242 FGFPWFAYWGQGTPTVYSSAK 262

Db 241 YYHYFDYWGQGTTLTVSSNK 261

RESULT 7

US-09-956-087-3

Sequence 3, Application US/09956087

Patent No. 6743908

GENERAL INFORMATION:

APPLICANT: FILPULA, DAVID

WANG, MAOLIANG

SHORR, ROBERT

WHITLOW, MARC

LEE, LIHSYNG S.

TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS

CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

STREET: 1100 NEW YORK AVE., NW, SUITE 600

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/956,087

FILING DATE: 20-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/069,821

FILING DATE: 1998-04-30

APPLICATION NUMBER: US 60/063,074

FILING DATE: 27-OCT-1997

APPLICATION NUMBER: US 60/050,472

FILING DATE: 23-JUN-1997

APPLICATION NUMBER: US 60/044,449

FILING DATE: 30-APR-1997

ATTORNEY/AGENT INFORMATION:

NAME: KIM, JUDITH U.

REGISTRATION NUMBER: 40,679

REFERENCE/DOCKET NUMBER: 0977.2280003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)371-2600

TELEFAX: (202)371-2540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 263 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-956-087-3

Query Match 37.0%; Score 870; DB 4; Length 263;

Best Local Similarity 68.2%; Pred. No. 5.2e-61;

Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;

QY 20 DIQTQSPSSLSASVGRVITCKASQ---DVGTSVAMYQKPKGKAPKLLIYWTSTRHTG 76

Db 1 DIQTQSPSSLSASVGRVITCKASQSLVSIINLAWYQKPKGKAPKLLIYAASSLESG 60

QY 77 VPSRFGSGSGCTDFTTISLQPEDFATYYCQYSLYR--SFGQGTKEIKRGGSGG-- 132

Db 61 VPSRFGSGSGCTDFTTISLQPEDFATYYCQYNSLPFWTGGQTKVEIKGTSGSGKP 120

QY 133 GSGGSGSGSEVQLVESGGVQVPGKSLRLSCASGFDFTTVMWVNRQAPKGLWIGEI 192

Db 121 GSGEGSTKGEVQLVESGGGLVQPGGSLRLSCAASGFTFSSYAMSVNRQAPKGLWVSVI 180

QY 193 --HPDSSTINVAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCA-----SLY 241

Db 181 SKGTDGGSTYYADSVKGRFTISRDNKNTLYLQWNSLRAEDTAVYYCARGXGXLSGX 240

QY 242 FGFPWFAYWGQGTPTVYSSAK 262

Db 241 YYHYFDYWGQGTTLTVSSNK 261

RESULT 8

US-09-135-121B-7

Sequence 7, Application US/09135121B

Patent No. 6534051

GENERAL INFORMATION:

APPLICANT: Dornburg, Ralph C.

TITLE OF INVENTION: CELL TYPE SPECIFIC GENE TRANSFER USING RETROVIRAL VECTORS

TITLE OF INVENTION: CONTAINING ANTIBODY-ENVELOPE FUSION PROTEINS AND WILD-TYPE

FILE REFERENCE: 97, 216-L

CURRENT APPLICATION NUMBER: US/09/135,121B

CURRENT FILING DATE: 1998-08-17

PRIOR APPLICATION NUMBER: US 08/933,616

PRIOR FILING DATE: 1997-08-28

PRIOR APPLICATION NUMBER: US 08/205,980

; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 07/979,619
; PRIOR FILING DATE: 1992-11-20
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 7
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Anti-Her2neu single chain antibody
US-09-135-121B-7

Query Match 35.7%; Score 838; DB 4; Length 332;
Best Local Similarity 48.5%; Pred. No. 2.4e-58;
Matches 180; Conservative 53; Mismatches 76; Indels 62; Gaps 10;

QY 6 IILFLVATATGVHS--DIQLTQSPSSLSASVGRVITITCKASODVGTSAVWYQKPKAP 63
DB 7 IFSLFLISASVIASRGDIVATQPKFASTSVGRISVTCKAS-DVGNVAWYQKPKQSP 65
QY 64 KLIYWTSTRTHTGVPSRFGSGSGTDFTTISLQPEDIAITYYCOQVSLYR-SFGQGTKV 122
DB 66 KPIIYSASYLYNGVDPRTGSGSGTDFSLISNVQSDDLAEYFCQYNTYPTFTFGGKTL 125
QY 123 EIKRGGSGSGSGSGSGSEVLVESGGVVGQPSRLSLSCSASGDFFTTYWMSWVRQAP 182
DB 126 EIKGSGSGSGSGSEG--KGEVLEESGGGLVQPKGSLKLSCAASGFTFTNTYAANWVRQAP 183
QY 183 GKLEWIGEHPSSITNYA-----PSLKDRFTISRDNKNTLFLQMSLRLPDTGYVFCFA 238
DB 184 GKLEWIVIRKSN--NYATYVDSVKDRFTISRDDSQALYLOANLKTEDTAAYYCV 241
QY 239 SLVFGEP--WFAYWGQGTPTVSSAKPTTTPAPRPTPTPTIASQPLSLRPEARPAAGG 296
DB 242 TSIDYDKVLFAWGQTTVTVSSA----- 266
QY 297 AVHTRGLDFALDKPLCYLLGILFIYGVILTALFLRVKFS-RSAEPPAYOQGNLYNEL 355
DB 267 -----DQQLCYLDAILFLYGLVLTLYCLRLKIQVRKAATISYEKSDG-VYTGL 314
QY 356 NLGRREYDVL 366
DB 315 STRNQETYL 325

RESULT 9
US-08-862-124-14
; Sequence 14, Application US/08862124
; Patent No. 6207153
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Maici, Pradip K.
; APPLICANT: Kaplan, Howard A.
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND
; TITLE OF INVENTION: DETECTION OF CANCERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster LLP
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,124

; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 31608-20001.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 304 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-862-124-14

Query Match 35.1%; Score 824; DB 3; Length 304;
Best Local Similarity 61.9%; Pred. No. 2.7e-57;
Matches 172; Conservative 29; Mismatches 51; Indels 26; Gaps 6;

QY 6 IILFLVATATGVHS--DIQLTQSPSSLSASVGRVITITCKASQDVGTSAVWYQKPKAPK 64
DB 10 IAVLAGPATVAQADIVLTQSPCTLSLSPGERATLSCLASQSVSSYLAWYQKPKQAPR 69
QY 65 LLIYWTSTRTHTGVPSRFGSGSGTDFTTISLQPEDIAITYYCOQY-----SLYRSFGQG 119
DB 70 LLIYGASTRATGMPDRFSGSGTDFTLTISRLEPEDFAVYCOQYSGSSPTQTITFGGG 129
QY 120 TKVEIKR-----GGSGSGSGSGSGSEVLVESGGVVGQPSRLSLSCSASGDFPT 171
DB 130 TKVEIKRTVAAPSVSGSGSGSGSGVQVLESVGGVVGQPSRLSLSCSASGDFPT 189
QY 172 TYWMSWVRQAPKGLWIGELIHPDSTINYPKDRFTISRDNKNTLFLQMSLRLPDT 231
DB 190 SPAMHWVRQALGKLEWAVISYDGTSTKYADSVKGRFTISRDTSKNTVYLKNSLTED 249
QY 232 TGVYFCA---SL-----YFGPFWFAYWGQGTPTVSS 260
DB 250 TAVYCARQDSLLGDYDHYGLD---VWGKGTITVSS 284

RESULT 10
US-08-488-113B-147
; Sequence 147, Application US/08488113B
; Patent No. 5744580
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,113B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/064,691
/ FILING DATE: 12-MAY-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/988,430
/ FILING DATE: 09-DEC-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McNicholas, Janet M.
/ REGISTRATION NUMBER: 32,918
/ REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/707-8889
/ TELEFAX: 312/707-9155
/ TELEX: 650 388-1248
/ INFORMATION FOR SEQ ID NO: 147:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 240 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-488-113B-147

Query Match 34.9%; Score 819.5; DB 1; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

QY 20 DIQTQSPSSLSASVGDRTVITCKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSSLSASVGDRTVITCKASQDVGTSAVYQOKPGKAPKLLIYRANRLESVPS 60

QY 80 RFSGSGSGTDFTTISLSQPEDIAITYCQYSLYR-SFGQGTKEIKRGSGSGSGSGG 138
Db 61 RFSGSGSGTDYTLTISLSQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGSGSGG 119

QY 139 SGSEVQLVESGGGVQPRSLRSLSCASGDFPTTVMWSVROAPKGLIEWGEIHPDST 198
Db 120 GGSEIQLVQSGGGLVKPGSVRAISCAASGYTFTNYGMNVVROAPKGLIEWGWINHTHGE 179

QY 199 INYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPW-FAYWGGGTPTV 257
Db 180 PTYADSFKGRFTSLDSDSKNTAYLIQINSRAEDTAVYFCTRR--GYDWFYDVWGGGTIVT 237

QY 258 VSS 260
Db 238 VSS 240

RESULT 11
US-08-477-484B-147
/ Sequence 147, Application US/08477484B
/ Patent No 5756659
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studnika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ NUMBER OF SEQUENCES: 169
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: McAndrews, Held & Malloy, Ltd.
/ STREET: 500 West Madison Street, 34th floor
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60661
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
```

```
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,484B
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/425,336
/ FILING DATE: 18-APR-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/064,691
/ FILING DATE: 12-MAY-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/988,430
/ FILING DATE: 09-DEC-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: McNicholas, Janet M.
/ REGISTRATION NUMBER: 32,918
/ REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/707-8889
/ TELEFAX: 312/707-9155
/ TELEX: 650 388-1248
/ INFORMATION FOR SEQ ID NO: 147:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 240 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-477-484B-147

Query Match 34.9%; Score 819.5; DB 1; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

QY 20 DIQTQSPSSLSASVGDRTVITCKASQDVGTSAVYQOKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSSLSASVGDRTVITCKASQDVGTSAVYQOKPGKAPKLLIYRANRLESVPS 60

QY 80 RFSGSGSGTDFTTISLSQPEDIAITYCQYSLYR-SFGQGTKEIKRGSGSGSGSGG 138
Db 61 RFSGSGSGTDYTLTISLSQYEDFGIYCCQYDESPWTFGGTKLEMK-GGGSGSGSGG 119

QY 139 SGSEVQLVESGGGVQPRSLRSLSCASGDFPTTVMWSVROAPKGLIEWGEIHPDST 198
Db 120 GGSEIQLVQSGGGLVKPGSVRAISCAASGYTFTNYGMNVVROAPKGLIEWGWINHTHGE 179

QY 199 INYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPW-FAYWGGGTPTV 257
Db 180 PTYADSFKGRFTSLDSDSKNTAYLIQINSRAEDTAVYFCTRR--GYDWFYDVWGGGTIVT 237

QY 258 VSS 260
Db 238 VSS 240

RESULT 12
US-08-646-360-147
/ Sequence 147, Application US/08646360
/ Patent No. 5837491
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studnika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ MEDIUM TYPE: Floppy disk
```

```

; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-646-360-147

Query Match 34.9%; Score 819.5; DB 2; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

Qy 20 DIQLTQSPSLASVGRVITTCASQDVGTSAVYQOQKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSLASVGRVITTCRASQDINSYLSWFQOQKPGKAPKLLIYRANRLESVPS 60
Qy 80 RFGSGSGTDYTLTISSQLQVEDFIYCCQYDSESPWTFGGTKLEMK-GGGSGSGSGG 138
Db 61 RFGSGSGTDYTLTISSQLQVEDFIYCCQYDSESPWTFGGTKLEMK-GGGSGSGSGG 119
Qy 139 SGSEVQLVESGGGVQPGFSLRLSCASGDFFTYTMMSWVRQAPGKGLVIGIHDPDSST 198
Db 120 GGSEIQLVQSGGGLVPGGSGVRISCAASGYFTFTNYGMNWRQAPGKGLWGWINTHTGE 179
Qy 199 INTAPSLKDRFTTISRDNAKNTLFLOMDSLRLPEDTCVYFCASLFGFPW-FAYWGQGTPT 257
Db 180 PTYADSPKGRFTTSLDSDSKNTAYLQINSRAEDTAVYFCTRR--GYDWFVDMVQGGTPT 237
Qy 258 VSS 260
Db 238 VSS 240
```

```

RESULT 13
US-08-839-765-147
; Sequence 147, Application US/08839765
; Patent No. 6146631
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/839,765
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-839-765-147

Query Match 34.9%; Score 819.5; DB 3; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;

Qy 20 DIQLTQSPSLASVGRVITTCASQDVGTSAVYQOQKPGKAPKLLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSLASVGRVITTCRASQDINSYLSWFQOQKPGKAPKLLIYRANRLESVPS 60
Qy 80 RFGSGSGTDYTLTISSQLQVEDFIYCCQYDSESPWTFGGTKLEMK-GGGSGSGSGG 138
Db 61 RFGSGSGTDYTLTISSQLQVEDFIYCCQYDSESPWTFGGTKLEMK-GGGSGSGSGG 119
Qy 139 SGSEVQLVESGGGVQPGFSLRLSCASGDFFTYTMMSWVRQAPGKGLVIGIHDPDSST 198
Db 120 GGSEIQLVQSGGGLVPGGSGVRISCAASGYFTFTNYGMNWRQAPGKGLWGWINTHTGE 179
```


TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 147:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-610-838-147

Query Match 34.9%; Score 819.5; DB 3; Length 240;
Best Local Similarity 65.8%; Pred. No. 4.5e-57;
Matches 160; Conservative 31; Mismatches 47; Indels 5; Gaps 4;
Qy 20 DIQLTSPSSLASVGRVITTCASQDVGTSTVAWYQOKPKAPKLIYWTSTRHTGVPS 79
Db 1 DIQMTQSPSSLASVGRVITTCASQDINSYLSWFQOKPKAPKTLIYRANLESGVPS 60
Qy 80 RFGSGSGTDFTFTISSLPEDIATYYCQOYSLYR-SFGQGTKVEIKRGGSGSGSGGG 138
Db 61 RFGSGSGTDYTLTSSLQVEDFGIYCCQYDESPWTFGGGTKLEMK-GGGSGSGGGGG 119
Qy 139 SGSEVOLVESGGGVQGRSLRLSCSASGDFTTYMMSWYRQAPGKLEWIGIHPDSST 198
Db 120 GSSEIQLVQSGGGLVPGGSVRISCAASGYFTNYGMNWRQAPGKLEWMGWINTHTGE 179
Qy 199 INYAPSLKDRFTISRDNKNTLFLQMDSLRPEDTGVYFCASLYFGFPW-FAYWGQGTPTV 257
Db 180 PTYADSFKGRFTSLDSDSKNTAYLIQINSRAEDTAVYFCTRR--GYDWYFDVWGQGTTVT 237
Qy 258 VSS 260
Db 238 VSS 240

Search completed: May 26, 2005, 02:16:51
Job time : 96 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2005, 02:12:59 ; Search time 337 Seconds
(without alignments)
439.724 Million cell updates/sec

Title: US-10-006-771b-2

Perfect score: 2350

Sequence: 1 MGWSCILFLVATATGVHSD.....LSTATKDYDALHMQALPFR 443

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2350	100.0	443	13	US-10-006-773-2
2	2350	100.0	443	13	US-10-006-771a-2
3	1228	52.3	634	16	US-10-416-011-2
4	1205.5	51.3	364	16	US-10-829-388-11
5	1205.5	51.3	370	16	US-10-829-388-1
6	1202.5	51.2	358	16	US-10-829-388-12
7	1202.5	51.2	363	16	US-10-829-388-2
8	1094	46.6	444	8	US-08-812-393A-2
9	1092	46.5	449	10	US-09-774-681-2
10	965	41.1	352	10	US-09-203-958A-2
11	910	38.7	631	15	US-10-120-198B-2
12	871	37.1	283	9	US-09-983-580-6
13	871	37.1	283	9	US-09-985-442-6

14	870	37.0	263	9	US-09-956-086-3	Sequence 3, Appli
15	870	37.0	263	9	US-09-956-087-3	Sequence 3, Appli
16	865.5	36.8	247	15	US-10-423-847-13	Sequence 13, Appl
17	865.5	36.8	247	17	US-10-831-063-13	Sequence 13, Appl
18	865.5	36.8	252	15	US-10-423-847-10	Sequence 10, Appl
19	865.5	36.8	252	17	US-10-831-063-10	Sequence 10, Appl
20	865.5	36.8	253	15	US-10-423-847-11	Sequence 11, Appl
21	865.5	36.8	253	17	US-10-831-063-11	Sequence 11, Appl
22	865.5	36.8	254	15	US-10-423-847-17	Sequence 17, Appl
23	865.5	36.8	254	17	US-10-831-063-17	Sequence 17, Appl
24	862.5	36.7	252	15	US-10-423-847-14	Sequence 14, Appl
25	862.5	36.7	252	17	US-10-831-063-14	Sequence 14, Appl
26	846	36.0	239	15	US-10-423-847-18	Sequence 18, Appl
27	846	36.0	239	17	US-10-831-063-18	Sequence 18, Appl
28	842.5	35.9	252	9	US-09-971-543-1	Sequence 1, Appli
29	838	35.7	332	14	US-10-211-488-7	Sequence 7, Appli
30	824	35.1	304	10	US-09-782-397-14	Sequence 14, Appl
31	824	35.1	304	15	US-10-651-453-14	Sequence 14, Appl
32	821	34.9	245	17	US-10-864-818-1	Sequence 1, Appli
33	819.5	34.9	240	14	US-10-127-890-147	Sequence 147, App
34	819.5	34.9	240	17	US-10-717-243-147	Sequence 147, App
35	808	34.4	491	13	US-10-011-125-2	Sequence 2, Appli
36	799.5	34.0	550	14	US-10-207-655-270	Sequence 270, App
37	780	33.2	266	14	US-10-207-655-260	Sequence 260, App
38	775	33.0	352	16	US-10-333-235A-54	Sequence 54, Appl
39	775	33.0	468	16	US-10-333-235A-56	Sequence 56, Appl
40	775	33.0	507	14	US-10-074-596-11	Sequence 11, Appl
41	775	33.0	565	16	US-10-333-235A-57	Sequence 57, Appl
42	774.5	33.0	374	14	US-10-335-394-15	Sequence 15, Appl
43	774	32.9	502	16	US-10-679-620-88	Sequence 88, Appl
44	772	32.9	456	15	US-10-634-862-11	Sequence 11, Appl
45	770.5	32.8	4852	15	US-10-412-406-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1

US-10-006-773-2
; Sequence 2, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Jungsans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-773-2

Query Match 100.0%; Score 2350; DB 13; Length 443;
Best Local Similarity 100.0%; Pred. No. 9.6e-150;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGWSCILFLVATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAWYQKPG	60
Db	1	MGWSCILFLVATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAWYQKPG	60
Qy	61	KAPKLLIYWTSTRTHTGVPRFSGSGTDFTTFTSSLPEDIIATYCCQYSLYRSFQGT	120
Db	61	KAPKLLIYWTSTRTHTGVPRFSGSGTDFTTFTSSLPEDIIATYCCQYSLYRSFQGT	120
Qy	121	KVEIKRGGSGSGSGSGSGSEVOLYVESGGVYQPGSLRLSCASGFDFTYVMSVWRQ	180
Db	121	KVEIKRGGSGSGSGSGSGSEVOLYVESGGVYQPGSLRLSCASGFDFTYVMSVWRQ	180

```

QY 181 APKGLEWGEIHPDSSINTVAPSLKDRFTISRDNKAKNTLFLQMSLRPDTGVVFCASL 240
Db 181 APKGLEWGEIHPDSSINTVAPSLKDRFTISRDNKAKNTLFLQMSLRPDTGVVFCASL 240
QY 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSLRPEAPPAAGGAVHT 300
Db 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSLRPEAPPAAGGAVHT 300
QY 301 RGLDPALDPKCLYLDGLIFGYVLTALFLRVKFSRSEAPPAYQOGONQLYNELNLGRR 360
Db 301 RGLDPALDPKCLYLDGLIFGYVLTALFLRVKFSRSEAPPAYQOGONQLYNELNLGRR 360
QY 361 BEYDVLDRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGL 420
Db 361 BEYDVLDRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGL 420
QY 421 YQGLSTATKOTYDALHMOALPPR 443
Db 421 YQGLSTATKOTYDALHMOALPPR 443

RESULT 2
US-10-006-771A-2
; Sequence 2, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250,090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-771A-2

Query Match 100.0%; Score 2350; DB 13; Length 443;
Best Local Similarity 100.0%; Pred. No. 9.6e-150;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMVYQQKPG 60
Db 1 MWSCIIILFLVATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMVYQQKPG 60
QY 61 KAPKLLIYWTSTRHTGVPSRSGSGGTDFTTISLQPEDIAITYCOOYSLYRSFGOGT 120
Db 61 KAPKLLIYWTSTRHTGVPSRSGSGGTDFTTISLQPEDIAITYCOOYSLYRSFGOGT 120
QY 121 KVEIKRGGSGGGSGGSEVQLVESGGGVQPGRLSLRSCSASGDFDTTVMWSWRQ 180
Db 121 KVEIKRGGSGGGSGGSEVQLVESGGGVQPGRLSLRSCSASGDFDTTVMWSWRQ 180
QY 181 APKGLEWGEIHPDSSINTVAPSLKDRFTISRDNKAKNTLFLQMSLRPDTGVVFCASL 240
Db 181 APKGLEWGEIHPDSSINTVAPSLKDRFTISRDNKAKNTLFLQMSLRPDTGVVFCASL 240
QY 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSLRPEAPPAAGGAVHT 300
Db 241 YFGPWFAYWGGTPTVTVSSAKPTTTPAPRPTTAPTASQPLSLRPEAPPAAGGAVHT 300
QY 301 RGLDPALDPKCLYLDGLIFGYVLTALFLRVKFSRSEAPPAYQOGONQLYNELNLGRR 360
Db 301 RGLDPALDPKCLYLDGLIFGYVLTALFLRVKFSRSEAPPAYQOGONQLYNELNLGRR 360
QY 361 BEYDVLDRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGL 420
Db 361 BEYDVLDRGRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGL 420

```

```

QY 421 YQGLSTATKOTYDALHMOALPPR 443
Db 421 YQGLSTATKOTYDALHMOALPPR 443

RESULT 3
US-10-416-011-2
; Sequence 2, Application US/10416011
; Publication No. US20040126363A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
; APPLICANT: Forman, Stephen
; APPLICANT: Raubitschek, Andrew
; TITLE OF INVENTION: CD19-specific redirected immune cells
; FILE REFERENCE: 1954-338
; CURRENT APPLICATION NUMBER: US/10/416,011
; CURRENT FILING DATE: 2003-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CD19R: zeta chimeric receptor
US-10-416-011-2

Query Match 52.3%; Score 1228; DB 16; Length 634;
Best Local Similarity 45.5%; Pred. No. 3.1e-74;
Matches 279; Conservative 53; Mismatches 91; Indels 190; Gaps 15;

QY 20 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMVYQQKPGKAPKLLIYWTSTRHTGVPS 79
Db 23 DIQMTQTSSLSASLGDRVTISCRASQDI SKLWYQQKPDGTVKLLIYHTRSLHSGVPS 82
QY 80 RFSGSGSGTDFTTISLQPEDIAITYCOQ--YSLYRSFGQGTKEIKRGGSGG--GSGS 136
Db 83 RFSGSGSGTDYSLTISNLEQEDIAITYFCQGNLTPTPTFGGCTKLEITGTSGSGKPGSGE 142
QY 137 GGSSEVOLVESGGGVQPGRLSLRSCSASGDFDTTVMWSWRQAPKGLWISGHIHPS 196
Db 143 GSTGKVLQSGPGLVAPQSLSVTCTVSGVSLFDYGVSWIRQPPRKGLEWLVIV--GS 201
QY 197 STINAPSLKDRFTISRDNKAKNTLFLQMSLRPDTGVVFCASLYF--GFPWFAYWGGT 254
Db 202 ETTYNGALKSRLLTIIDNKSQVFLKNSLQTDITAIYCKKHYGGSVAMDIWGGT 261
QY 255 PVTSSAKPTTTPAPR--PPTPAPTASQP-----LSLRPEA----- 289
Db 262 SVTVSSVEPKSSDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH 321
QY 290 ----- 291
Db 322 EDPEVKENWYVDGVEVHNNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNKAL 381
QY 292 PA-----AGG-----AVHT-----RGL---DFALD-----PK 310
Db 382 PAPEKTIKAGQPRPEQVTVLPSPRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 441
QY 311 LCYLL-----DGILFIY----- 322
Db 442 NNYKTPVPLDSDGSFFLYSKLTVDKSRWQQGNVPSCSVMHEALHNNHYTKSLSLSPGKM 501
QY 323 -----GVIL---TALFLRVKFSRSEAPPAYQOGONQLYNELNLGRRREYDVLDRR 370
Db 502 ALIVLGGVAGLLFIGLGIFPRVKFSRSEADAPAYQQGONQLYNELNLGRRREYDVLDRR 561
QY 371 GRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGLYQGLSTATKD 430
Db 562 GRDPEMGKPRKPNQEGLYNELQDKMAEYSEIGMKGERRRGKGDGLYQGLSTATKD 621
QY 431 TYDALHMOALPPR 443

```

```

Db      622 TYDALHMOALPPR 634

RESULT 4
US-10-829-388-11
; Sequence 11, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric sequence from multiple species
US-10-829-388-11

Query Match      51.3%; Score 1205.5; DB 16; Length 364;
Best Local Similarity 94.3%; Pred. No. 5.4e-73;
Matches 230; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

Qy      16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHT 75
Db      122 GGSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHT 181
Qy      76 GVPSRFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKRGSGSGSG 135
Db      182 GVPSRFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKR----- 232
Qy      136 SGGSGSEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPD 195
Db      233 LEGGSEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPD 292
Qy      196 SSTINYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 255
Db      293 SSTINYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 352
Qy      256 VTVS 259
Db      353 VTVS 356

RESULT 5
US-10-829-388-1
; Sequence 1, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1

Query Match      51.3%; Score 1205.5; DB 16; Length 364;
Best Local Similarity 94.3%; Pred. No. 5.4e-73;
Matches 230; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

Qy      16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHT 75
Db      122 GGSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHT 181
Qy      76 GVPSRFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKRGSGSGSG 135
Db      182 GVPSRFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKR----- 232
Qy      136 SGGSGSEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPD 195
Db      233 LEGGSEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPD 292
Qy      196 SSTINYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 255
Db      293 SSTINYAPSLKDRFTISRDNKNTLFLQMSLRPEDTGVYFCASLYFGFPWPAYWGQGP 352
Qy      256 VTVS 259
Db      353 VTVS 356

RESULT 6
US-10-829-388-12
; Sequence 12, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-0006US1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIOR FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric sequence from multiple species
US-10-829-388-12

Query Match      51.2%; Score 1202.5; DB 16; Length 358;
Best Local Similarity 95.0%; Pred. No. 8.5e-73;
Matches 228; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

Qy      20 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHTGVPS 79
Db      2 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSVAMYYQQKPGKAPKLLIYWTSTRHTGVPS 61
Qy      80 RFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKRGSGSGSGSGS 139
Db      62 RFGSGSGTDFTFTISSLPEDATYCCQYSLYRSFGQGTKEIKR-----GGQ 112
Qy      140 GSEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPDSTI 199
Db      113 FMEVLVESGGGVQGRSLRLSCSASGDFTTYMNSWVRQAPGKGLEWIGIHPDSTI 172

```

```

QY 200 NYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 259
DB 173 NYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 232

RESULT 7
US-10-829-388-2
; Sequence 2, Application US/10829388
; Publication No. US20050003403A1
; GENERAL INFORMATION:
; APPLICANT: ROSSI, EDMUND A.
; APPLICANT: CHANG, CHIEN HSING
; APPLICANT: MCBRIDE, WILLIAM J.
; TITLE OF INVENTION: POLYVALENT PROTEIN COMPLEX
; FILE REFERENCE: 41133-000605U1
; CURRENT APPLICATION NUMBER: US/10/829,388
; PRIORITY FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: 60/464,532
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: 60/525,391
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 363
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric sequence from multiple species
US-10-829-388-2

Query Match 51.2%; Score 1202.5; DB 16; Length 363;
Best Local Similarity 95.0%; Pred. No. 8.6e-73;
Matches 228; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 20 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSMVYQKPGKAPKLLIYWTSTRHTGVPS 79
DB 7 DIQLTQSPSSLSASVGDRTVITCKASQDVGTSMVYQKPGKAPKLLIYWTSTRHTGVPS 66

QY 80 RFGSGSGTDTFTTSSLPEDIAITYCQXSLYRSFGQTKVEIKRGGSGSGSGGS 139
DB 67 RFGSGSGTDTFTTSSLPEDIAITYCQXSLYRSFGQTKVEIKRG-----GGQ 117

QY 140 GSEVLVESGGVQVQGRSLRLSCSAGDFDTTYMWSVRQAPGKLEWIEIHPDSSTI 199
DB 118 FMEVLVESGGVQVQGRSLRLSCSAGDFDTTYMWSVRQAPGKLEWIEIHPDSSTI 177

QY 200 NYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 259
DB 178 NYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCASLYFGPPWYAYWGQGTPTVTS 237

RESULT 8
US-08-812-393A-2
; Sequence 2, Application US/08812393A
; Publication No. US20010007152A1
; GENERAL INFORMATION:
; APPLICANT: SHERMAN, Linda A.
; APPLICANT: LUSTGARTEN, Joseph
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING
; TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

```

; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812.393A
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 31333-20001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-812-393A-2

Query Match 46.6%; Score 1094; DB 8; Length 444;
Best Local Similarity 54.7%; Pred. No. 2.1e-65;
Matches 248; Conservative 35; Mismatches 106; Indels 64; Gaps 12;

QY 22 QLQTSPLSASVGDRTVITCKASQDVGTSMVYQKPGKAPKLLIYWTSTRHTGVPSR 80
DB 25 QVQSPASLVLEQENAELOCSFS--IFTNQVQVQFQPGGRLVSLLY-----NPSG 74

QY 81 FSGSGSGTDTFTTSSLPEDIAITYCQOYS-----LYRSFGQTKVEIKGG 128
DB 75 TKQSGRLTSTTVIKERRSSLSHSSQITDSGTIYLCASNSGGSNAKLTFGKGTKLVSXSGG 134

QY 129 SGSGSGSGSGSGSVQLVES--GGGVQVQGRSLRLSCSAGDFDTTYMWSVRQAPGKGLE 187
DB 135 GSGSGSGSGSGSGSEAAVTQSPRNKAVTGGKVTLSCTQNNHNMV---WTRQDTGHLR 191

QY 188 WIGIHPDSSTINTYAPSLKDRFTISRDNKNTLFLQMSLRPDTGVYFCASLYFGPPW 246
DB 192 LIHYSYAGST--BKGDIPDGYKASRPSQENFSLILELAT--PSQTSVYFCASGETGTNE 247

QY 247 PAYWQGTPTVTVSS-----AKPTTTPAPRPTTAPTIASQPLSLRPEAA 290
DB 248 RLFFGHGHTKLSVLTNSIMYFHFVFLPAKPTTTTAPRPTTAPTIASQPLSLRPS 307

QY 291 RPAAGGAVHTRGLDFALDPKLCYLLDGLTFYGVILTALFLRVKFSRSEAPPAYQOQNG 350
DB 308 R-----DPKLCYLLDGLTFYGVILTALFLRVKFSRSEADAPAYQOQNG 351

QY 351 LYNELNLGRREYDVLDRGRDPEMGKPKRKNPQEGLYNELQDKMAEAYSEIGMKGE 410
DB 352 LYNELNLGRREYDVLDRGRDPEMGKPKRKNPQEGLYNELQDKMAEAYSEIGMKGE 411

QY 411 RRRKGHDGLYQGLSTATKOTYDALHMQALPPR 443
DB 412 RRRKGHDGLYQGLSTATKOTYDALHMQALPPR 444

RESULT 9
US-09-774-681-2
; Sequence 2, Application US/09774681
; Publication No. US20030208780A1
; GENERAL INFORMATION:
; APPLICANT: Sunol Molecular Corporation
; APPLICANT: Sherman, Linda
; APPLICANT: Lustgarten, Joseph

```



```

; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
; FILE REFERENCE: 31333-20001.01
; CURRENT APPLICATION NUMBER: US/09/774,681
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 08/812,393
; PRIOR FILING DATE: 1997-03-05
; PRIOR APPLICATION NUMBER: US 60/012,845
; PRIOR FILING DATE: 1996-03-05
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid derivative of effective T cell
; OTHER INFORMATION: receptor
US-09-774-681-2

Query Match 46.5%; Score 1092; DB 10; Length 449;
Best Local Similarity 54.5%; Pred. No. 2.9e-65;
Matches 247; Conservative 36; Mismatches 106; Indels 64; Gaps 12;

Qy 22 QLTQSPSSLSASVGDRTVITCKASQDVGTSS-VAVYQKPGKAPKLLIYWTSTRHTGVPSSR 80
Db 25 QVQSPASLVQEGENAELOCSF--IFTNQVWFYQRPGRGLVSLIY-----NPSG 74
Qy 81 FSGSGSGTDTFT-----FTISLQPEDIATYCCQYS---LYRSFGOGTKVEIKRGG 128
Db 75 TKSGRLTSTTVIKERRSLHSSQITDSGYLCASSNGGSNAKLTFGKGTKLVSXSGG 134
Qy 129 SGSGSGSGSGSGSEVQLVES--GGVQVQPGRLSLRLSCSASGDFDFTYMWVWVROAPGKGL 187
Db 135 GSGSGSGSGSGSGSEAAVTQPRNKVAVTGKVLSCNQTNHNNY---WYRQDTGHGLR 191
Qy 188 WIGEHPDSTINYPALSKDRFTISRDNAXN-TLFLQMDSLRPEDTGVYFCASLYGFPW 246
Db 192 LIHYSYAGST--EKGDIPGKYASRPSQENFSLIVELGT--PSQTSVYFCASGETGTNE 247
Qy 247 FAYWQGTPTVTVSS-----AKPTTTPAPRPPTPAPTIASQPLSLRPEAA 290
Db 248 RLFPFGHGTKLSVLTNSIMYFHSHPVFLPAKPTTTPAPRPPTPAPTIASQPLSRPSS 307
Qy 291 RPAAGGAVHTRGLDFALDKPLCYLLDGLIFTYGVILTALFLRVKFSRSAPPAQOQONQ 350
Db 308 R-----DPLCYLLDGLIFTYGVILTALFLRVKFSRSADAPAYQOQONQ 351
Qy 351 LYNELNLGRREYVDLKRGRDPEMGKPRRKNPQEGLYNELQKQKVAEYSEIGMKGE 410
Db 352 LYNELNLGRREYVDLKRGRDPEMGKPRRKNPQEGLYNELQKQKVAEYSEIGMKGE 411
Qy 411 RRRKGHDGLYQGLSTATKTYDALHQAIPPR 443
Db 412 RRRKGHDGLYQGLSTATKTYDALHQAIPPR 444

RESULT 10
US-09-203-958A-2
; Sequence 2, Application US/09203958A
; Publication No. US20030039641A1
; GENERAL INFORMATION:
; APPLICANT: KELLER, Tibor
; APPLICANT: GOLDSTEIN, Joel
; APPLICANT: GRAZIANO, Robert
; APPLICANT: DEO, Yashwant M.
; TITLE OF INVENTION: CELLS EXPRESSING ANTI-FC RECEPTOR
; TITLE OF INVENTION: BINDING COMPONENTS
; FILE REFERENCE: MXI-099CPA
; CURRENT APPLICATION NUMBER: US/09/203,958A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: 60/067232
; PRIOR FILING DATE: 1997-12-02

```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-203-958A-2

Query Match 41.1%; Score 965; DB 10; Length 352;
Best Local Similarity 75.3%; Pred. No. 7.4e-57;
Matches 195; Conservative 13; Mismatches 43; Indels 8; Gaps 3;

Qy 12 ATATGVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSS-----VAVYQKPGKAPKL 65
Db 30 AGAQPASDIQLTQSPSSLSASVGDRTVITCKSSQSVLYSSNQKNLAWYQKPGKAPKL 89
Qy 66 LIYWTSTRHTGVPSSRFSFGSGSGTDTFTTISLQPEDIATYCCQYSLYRSFGOGTKVEIK 125
Db 90 LIYWASTRESGVPSRFSFGSGSGTDTFTTISLQPEDIATYCHQYLSSWTFGOGTKVEIK 149
Qy 126 -RGSGSGSGSGSGSEVQLVESGGVQVQPGRLSLRLSCSASGDFDFTYMWVWVROAPGK 184
Db 150 SSGSGSGSGSGSGSEVQLVESGGVQVQPGRLSLRLSCSSGFIIFSDNYWYVROAPGK 209
Qy 185 GLEWIGEHPDSTINYPALSKDRFTISRDNAXNTLFLQMDSLRPEDTGVYFCASLYGFG 244
Db 210 GLEWVATISDGGSYTYPPDSVKGRFTISRDNAXNTLFLQMDSLRPEDTGVYFCARGYRY 269
Qy 245 PW-FAYWQGTPTVTVSSAK 262
Db 270 EGAMDYWGQGTPTVTVSSPR 288

RESULT 11
US-10-120-198B-2
; Sequence 2, Application US/10120198B
; Publication No. US20030215427A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
; TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
; FILE REFERENCE: 1954-337
; CURRENT APPLICATION NUMBER: US/10/120,198B
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/282,859
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: mouse-human chimera
US-10-120-198B-2

Query Match 38.7%; Score 910; DB 15; Length 631;
Best Local Similarity 44.4%; Pred. No. 6.8e-53;
Matches 228; Conservative 47; Mismatches 118; Indels 120; Gaps 18;

Qy 16 GVHSDIQLTQSPSSLSASVGDRTVITCKASQDVGTSSVAVYQKPGKAPKLLIYWTSTRHT 75
Db 154 GGSDIQLTQSPSSLSASVGDRTVITCKANEDINNRLAWYQTPGNSPRLLISGATNLVT 213
Qy 76 GVPSRFSFGSGSGTDTFTTISLQPEDIATYCCQY--SLYRSFGOGTKVEIKRSGSGSGS 134
Db 214 GVPSRFSFGSGSGDKYTLTITSLQAEFPATYCCQYWSPTFTFGSGTELEIKVPEKSDKT 273
Qy 135 GSGSGSGSEVQLVESGGV-----QGRSLRLSCSASGDFDFTYMWVWVROAPGKLEWI 189
Db 274 HTCCPPCAPELL--GGPSVFLFPKPKDMLMISRTP---EVTGVVDVSHEDPEVKFNW- 327

```


Db	181	SKTGGSTYYADSVKGRFTISRDNKNLTLYLQMSLRADTAIVYCARGXKSLSGXY	240
Qy	242	FGFPMFAYWGQGTPVTVSSAK	262
Db	241	YYYHYFDYWGGTLTVSSNK	261
 RESULT 15 US-09-956-087-3 ; Sequence 3, Application US/09956087 ; Patent No. US20020161201A1 ; GENERAL INFORMATION: ; APPLICANT: FILPULA, DAVID ; WANG, MAOLIANG ; SHORR, ROBERT ; WHITLOW, MARC ; LEE, LIHSYNG S. ; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS ; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF ; NUMBER OF SEQUENCES: 33 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C. ; STREET: 1100 NEW YORK AVE., NW, SUITE 600 ; CITY: WASHINGTON ; STATE: DC ; COUNTRY: USA ; ZIP: 20005 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: Patentin Release #1.0, Version #1.30 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/09/956,087 ; FILING DATE: 20-Sep-2001 ; CLASSIFICATION: <Unknown> ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: 09/069,821 ; FILING DATE: 1998-04-30 ; APPLICATION NUMBER: US 60/063,074 ; FILING DATE: 27-OCT-1997 ; APPLICATION NUMBER: US 60/050,472 ; FILING DATE: 23-JUN-1997 ; APPLICATION NUMBER: US 60/044,449 ; FILING DATE: 30-APR-1997 ; ATTORNEY/AGENT INFORMATION: ; NAME: KIM, JUDITH U. ; REGISTRATION NUMBER: 40,679 ; REFERENCE/DOCKET NUMBER: 0977.2280003 ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: (202)371-2600 ; TELEFAX: (202)371-2540 ; INFORMATION FOR SEQ ID NO: 3: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 263 amino acids ; TYPE: amino acid ; STRANDEDNESS: single ; TOPOLOGY: peptide ; MOLECULE TYPE: peptide ; SEQUENCE DESCRIPTION: SEQ ID NO: 3: US-09-956-087-3 Query Match 37.0%; Score 870; DB 9; Length 263; Best Local Similarity 68.2%; Pred. No. 1.3e-50; Matches 178; Conservative 23; Mismatches 42; Indels 18; Gaps 5;			
Qy	20	DIQLTQSPSSLASGVDRVTITCKAQ---	DVGTSVANVYQQKPGKAPKLLIYTSTRHTG 76
Db	1	DIQMTQSPSSLASGVDRVTITCRASQSLVSISNYLAWYQQKPGKAPKLLIYAASSLESG	60
Qy	77	VPSRFSGSGGTFTFTISSLPEDATATYCOOYSLYR--SFQGTQVKIKRGSGSG--	132

Db 61 VPSRFGSGSGTDFTLTISLQPEDPATYCCOYNLSLPWTFGQGTKVEIKGSTSGSKP 120
QY 133 GSGSGSGSEVOLVESGGGVQGRSLRLSLCSASGDFDTTYWMSWVRQAPGKGLEWIGEI 192
Db 121 GSGEGSTKGEVOLVESGGGLVQPGGSLRLSLCAASGTFSSYAMSWVRQAPGKGLEWVSVI 180
QY 193 --HPDSSTINVAPSLKDRFTISRDNNAKNTLFLQMSLRPEDTGVYFCA-----SLY 241
Db 181 SGTDDGGSTYADSVKGRFTISRDNKNTLYLQMSLRAEDTAVYYCARGRXGSLSGXY 240
QY 242 FGFPMFAYWGCTPVTVSSAK 262
Db 241 YYYHYFDYWGQGLVTVSSNK 261

Search completed: May 26, 2005, 02:44:38
Job time : 338 secs